NUNC COGNOSCO EX PARTE

TRENT UNIVERSITY
LIBRARY

PRESENTED BY

Mrs. I. Hughson

4/6

Murray G. Hughson.

Digitized by the Internet Archive
in 2019 with funding from
Kahle/Austin Foundation

https://archive.org/details/principlesofecon0000garv

PRINCIPLES OF ECONOMICS

BY

FREDERIC B. GARVER, PH. D.
PROFESSOR OF ECONOMICS, UNIVERSITY OF MINNESOTA

AND

ALVIN HARVEY HANSEN, PH. D.
PROFESSOR OF ECONOMICS, UNIVERSITY OF MINNESOTA

GINN AND COMPANY
BOSTON · NEW YORK · CHICAGO · LONDON
ATLANTA · DALLAS · COLUMBUS · SAN FRANCISCO

HB171.5.G25

COPYRIGHT, 1928, BY FREDERIC B. GARVER AND ALVIN HARVEY HANSEN
ALL RIGHTS RESERVED

PRINTED IN THE UNITED STATES OF AMERICA

434.2

28-27671

The Athenæum Press
GINN AND COMPANY · PROPRIETORS · BOSTON · U.S.A.

PREFACE

The authors of this text believe that a thorough grasp of the principles of economics precedes an understanding of economic life. This holds both for the student of social problems and for the business man. The emphasis of this book is therefore on the fundamentals of production, price, and distribution.

But in presenting these principles to the university or college student the instructor is confronted by a twofold difficulty. A rigidly formal exposition of the theories and laws of economics, though it probably approaches most nearly the scientific ideal, is likely to leave the student with the feeling that the subject is highly abstract and that it has little relation to concrete facts as he encounters them in his limited contacts with economic life. On the other hand, a course that plunges the student into masses of statistical and descriptive data without sufficient explanation of the relations existing between these facts, and without providing him with a method for analyzing them, is equally unsatisfactory. The authors of this text have attempted to avoid these extremes. The principles have been kept constantly in the foreground, but throughout the book these principles have been applied to the interpretation of the problems of contemporary economic life.

Because of this method of continuous application of principles to problems, we believe that instructors who use this book will find an abundance of materials for classroom discussion, and will be able to elicit the interest of the student in the application of economic analysis to current economic problems. Satisfactory results have been obtained by the method here employed in the "Principles of Economics" course at the University of Minnesota, where a preliminary edition of this text has been in use for two years.

17663

The amount of space allotted in this book to the analysis of cost and price is somewhat greater than is ordinarily assigned to this subject. It is our belief that this subject is fundamental to a beginning course which is intended, in part at least, as a foundation for more detailed studies in various fields of economics and business administration. But in the chapters on this subject the actualities of economic life are constantly held in mind; and in turn the principles of overhead cost, joint demand, opportunity cost, increasing and decreasing cost are used as tools of analysis in the more descriptive chapters of the book. These doctrines of cost have not been set up as rules of thumb, which the student is expected to believe are applicable to real life without taking into account conditions not provided for in the theories.

The effects of institutional influences and of dynamic changes upon the more rigid conclusions of economic theory have not been minimized. We have endeavored to emphasize the dynamic phases of economic life not merely in the chapters on money and prices, institutional aspects of distribution, and international economic relations, but throughout the book.

We have made extensive use of illustrative diagrams and statistical charts in the belief that charts are much more significant to the student than tables of figures both in the analysis of theoretical principles and in the interpretation of historical data.

It is perhaps unnecessary to say that we make no claim to having made contributions to economic theory in this book.

We wish to acknowledge our indebtedness to the instructors in the course in "Principles of Economics" at the University of Minnesota for many helpful suggestions. We are also under obligation to Professor John D. Black of Harvard University, and to our colleagues, Professors R. S. Vaile and B. D. Mudgett and Dr. A. G. Black, for reading and criticizing special chapters.

F. B. GARVER
ALVIN H. HANSEN

MINNEAPOLIS, MINNESOTA

CONTENTS

BOOK I. PRODUCTION

CHAPTER PAGE

I. INTRODUCTION 3

The subject matter of economics, 3 · Economics as a science, 7.

II. PRODUCTION 15

Wants and production, 15 · Classification of productive activities, 17 · The source of utilities, 21 · Economic utilities, 21 · Production in the modern economic order, 22 · Control over utilities, 25 · Private income versus social income, 27.

III. THE REGULATION OF PRODUCTION 30

The individualistic order, 30 · The law of supply and demand, 31 · Assumptions of the individualistic philosophy, 31 · Public regulation, 33 · Unregulated private capitalism versus regulated capitalism, 34.

IV. THE ORGANIZATION OF PRODUCTION 40

The function of the entrepreneur, 40 · Forms of entrepreneurial organization, 43 · Individual proprietorship, 43 · Partnership, 44 · Business corporation, 45 · Coöperative society, 47 · Political units as entrepreneurs, 48.

V. COMBINATION OF THE AGENTS OF PRODUCTION 52

The background of social institutions, 52 · The factors of production, 53 · The relative scarcity of the factors, 55 · The law of diminishing returns, 56 · Agents of production, 60 · The best combination of the agents of production, 60 · Average product and marginal product, 63 · Cost per unit of output, 64 · Lowest unit cost for marginal entrepreneurs, 65 · Lowest unit cost for superior entrepreneurs, 67 · Large-scale production, 72.

VI. MARKETS AND THE MARKETING FUNCTION IN PRODUCTION 80

Definition and classification of markets, 80 · The variability of price-making forces, 82 · Marketing functions, 86 · Marketing and specialization, 90.

BOOK II. VALUE

VII. THE PRICE-MAKING PROCESS: DEMAND AND SUPPLY . . 97

Price and value, 97 · The analysis of price, 98 · Market price, 100 · Demand and demand price, 101 · The individual-demand schedule, 103 · Demand curves, 105 · Collective demand, 107 · Characteristics of demand schedules, 108 · Changes in demand, 112 · Supply and supply schedules, 114 · The determination of supply, 120 · Changes in supply, 121.

CHAPTER PAGE

VIII. THE PRICE-MAKING PROCESS: MARKET PRICE 123

Reserved and unreserved supply, 123 · Price without reservation of supply, 123 · The significance of supply prices, 129 · Market price with reservation of supply, 129 · Effects of changes in demand; supply constant, 130 · Effects of changes in supply; demand constant, 131 · Effects of changes in both demand and supply, 131 · Marginal buyers and sellers, 132 · Equilibrium prices and actual prices, 134.

IX. CONSUMERS' DEMAND AND UTILITY 140

The analysis of individual and collective demand, 140 · Utility, 143 · Utility and scarcity, 145 · The principle of diminishing utility, 146 · Marginal utility, 149 · Utility and the individual-demand schedule, 151 · Marginal utility and market price, 157 · Some practical applications of the marginal-utility theory, 158.

X. COST OF PRODUCTION AND NORMAL PRICE 162

How supply is determined, 162 · Cost of production and price, 163 · Normal price, 165 · Prime and supplementary costs, 167.

XI. DIFFERENTIAL COSTS AND NORMAL PRICE 174

Competition and prices, 174 · Entrepreneurs' differential costs and normal price, 176 · The permanence of entrepreneurs' cost differentials, 183.

XII. INTERDEPENDENT PRICES 190

The principal types of interdependent prices, 190 · Rival supply and price, 193 · Joint demand, 197 · Rival demand and price, 202 · Joint supply and joint cost, 205 · Overhead costs not joint costs, 206 · Summary of the theory of normal value, 208.

XIII. CONSTANT COST 211

Immediate effects of a change in demand, 211 · Effects of changes in demand for longer periods, 212 · Constant cost, 214 · Conditions under which constant cost prevails, 215 · Demand and price under conditions of constant cost, 217 · Constant cost and the shifting of costs, 217.

XIV. INCREASING COST 220

The meaning of increasing cost, 220 · Conditions leading to increasing cost, 222 · Diminishing returns and increasing expenses of production, 224 · Entrepreneurs' differential expenses and increasing cost, 226 · Forces counteracting increasing cost, 227 · Increasing cost and the incidence of new unit costs, 230.

XV. DECREASING COST 235

The meaning of decreasing cost, 235 · Causes of decreasing cost, 237 · Historical trend of cost and decreasing cost, 244 · Decreasing cost and restricted competition, 246 · Decreasing cost and the incidence of new unit costs, 248 · Increasing and decreasing cost compared, 249.

CHAPTER PAGE

XVI. MONOPOLY PRICE 253

The meaning of monopoly, 253 · The theory of monopoly price, 255 · The causes of monopoly, 257 · Limited competition, 260 · Limitations on monopoly price, 262.

XVII. MONOPOLY AND RESTRICTED COMPETITION 265

Forms of combination, 265 · The Sherman Anti-Trust Law, 268 · The "rule of reason," 269 · Efficiency versus unfair practices, 270 · Trade associations, 272 · The Federal Trade Commission and the Clayton Act, 272 · Illustrations of monopoly based on unfair practices, 273 · Quasi-monopoly based on efficiency, 277 · Monopoly based on patent rights, 278 · International Kartells, 279 · International concerns, 280 · Price policies, 281 · The modern organized economy, 283.

XVIII. PUBLIC REGULATION OF MONOPOLY PRICES; PUBLIC UTILITIES 287

The nature of a public utility, 287 · Economic characteristics of public utilities, 289 · Regulation of public-utility earnings, 291 · Market value of utility property, 292 · Original cost, 293 · Cost of reproduction, 297 · Specific-rate principles, 299 · How railroad rates are actually made, 301 · The Transportation Act of 1920, 301 · Federal control over security issues, 302 · Railroad consolidation, 302.

XIX. SPECULATION AND PRICES 305

Production for the future, 305 · Methods of meeting business risks, 306 · Speculation and risk-taking, 309 · Effect of speculation on prices, 311.

BOOK III. MONEY AND PRICES

XX. MONEY AND CREDIT 317

The function of money, 317 · Money and circulating media, 318 · Kinds of money in the United States, 318 · The standard of value, 322 · Deposit currency, 327 · Credit instruments, 328 · The volume of deposit currency, 329 · The federal reserve system, 331 · Our monetary system, 333.

XXI. PRICE FLUCTUATIONS 336

Relative prices, or index numbers, 336 · The making of index numbers, 338 · History of prices, 339 · Effect of price fluctuations, 341.

XXII. THE THEORY OF MONEY AND PRICES 349

The value of money, 349 · The equation of exchange, 351 · The quantity theory, 355 · The income theory, 365 · Stabilization of the value of money, 367 · Stabilization of the value of gold, 369.

CHAPTER PAGE

XXIII. BUSINESS CYCLES 373

Industrial fluctuations, 373 · Various phases of the business cycle, 375 · Rainfall and business cycles, 382 · Overproduction and underconsumption, 383 · Money and credit, 387 · Self-generating cycles, 388 · Price fluctuations and the business cycle, 390 · Progress and the business cycle, 391 · The business cycle an unsolved problem, 393 · Stabilization, 394 · Value of elimination of the business cycle, 399.

BOOK IV. THE DISTRIBUTION OF WEALTH AND INCOME

PART I. FUNCTIONAL DISTRIBUTION

XXIV. WAGES 403

The meaning of distribution, 403 · The importance of distribution, 404 · The earnings of labor, 406 · Wages the price of labor, 407 · The marginal productivity of labor, 409 · The principle of substitution, 413 · The supply of labor, 415 · Improvements in technique and wages, 423 · Principles of wage-adjustments, 427.

XXV. WAGE DIFFERENCES 429

Labor-income groups, 429 · The proximate cause of differences in wages, 433 · Fundamental causes of differences in wages, 434 · Differences in wages between different places, 437 · The wages of women, 440 · Minimum-wage legislation, 440.

XXVI. THE RENT OF LAND 443

Land as a factor in production, 443 · The meaning of "rent," 444 · The demand for land, 446 · The amount of rent, 449 · Rent and differences in land, 450 · The generalized, or social, theory of rent, 454 · Agricultural rent, 456 · Urban rent, 457 · The unearned increment, 459.

XXVII. INTEREST 464

Nature and kinds of capital, 464 · Kinds of producers' capital, 466 · The return to capital goods, 468 · The superior productivity of capital, 470 · Time preference, 472 · The motives for saving, 475 · The supply of savings, 478 · The demand for savings, 480 · The capitalization of incomes, 484 · Types of interest, 486 · The loan market, 487 · Interest rates and the supply of money, 489.

XXVIII. PROFITS 493

Business profits and economic profits, 493 · Pure profits as a share in distribution, 495 · Profits from the earnings of rare natural ability, 498 · Uncertainty in economic life, 499 · The entrepreneur as risk-taker, 501 · Risk-taking and profits, 502 · The real cost of risk-taking, 504 · Natural limitations on competition, 505 · Gains of monopoly and of exploitation, 506 · Summary and conclusion, 507.

PART II. INSTITUTIONAL ASPECTS OF DISTRIBUTION

CHAPTER PAGE

XXIX. TRADE UNIONISM AND LABOR LEGISLATION 509

The origin of trade unions, 509 · How trade unions affect wages, 511 · The growth of unionism in the United States, 516 · Employers' associations, 520 · Trade-union demands, 520 · The methods of unionism, 523 · Collective bargaining, 524 · Unionism in the modern industrial order, 526 · The legal status of unionism, 528 · Government intervention in labor disputes, 532 · Minimum-wage legislation, 535 · Hours of work, 536 · Social insurance, 537 · Economic principles of labor legislation, 538 · Panaceas and the labor problem, 541.

XXX. DYNAMIC ASPECTS OF DISTRIBUTION 544

The wage-earning class; wages, 544 · The effect of productivity on wages, 546 · The effect of price fluctuations on wages, 548 · The effect of immigration on wages, 553 · The effect of unionism on wages, 555 · The farming class; land value and land rent, 555 · The business class; profits, 563 · The investing class; interest rates, 568 · Other factors affecting the rate of interest, 573.

XXXI. INEQUALITIES OF PROPERTY-OWNERSHIP AND INCOME 576

Distribution of wealth in Massachusetts and the United States, 577 · Distribution of wealth in two Minnesota counties, 584 · Distribution of income, 586 · Equalitarian distribution considered, 590 · Production versus distribution, 592 · Permanency of relative distribution of income, 593 · Property institutions and social utility, 597 · Functional distribution, 598 · Functional and personal distribution contrasted, 599.

XXXII. SOCIALISM 602

The extent of the socialist movement, 602 · Origin and development of socialism, 603 · Utopian socialism, 604 · Marxian socialism, 605 · Evolutionary socialism, 608 · Syndicalism, 609 · Guild socialism, 610 · Communism, 611 · Socialism and communism contrasted, 613 · Socialism and capitalism, 614 · Difficulties facing socialism, 615 · Waste in the present order, 618 · Value and distribution in a socialist order, 619.

XXXIII. SOME ECONOMIC ASPECTS OF TAXATION 622

Government finance and economic principles, 622 · The nature of a tax, 624 · Nontax revenue, 625 · Kinds of taxes, 626 · The incidence of taxes, 627 · The capitalization of taxes, 631 · The amount of taxation, 633 · Increase in taxation, 635 · Tests of a good tax system, 640 · The distribution of taxation, 640 · Progressive taxation versus proportional taxation, 644.

BOOK V. INTERNATIONAL ECONOMIC RELATIONS

CHAPTER PAGE

XXXIV. FOREIGN EXCHANGE 649

The meaning of foreign exchange, 649 · Forms of foreign exchange, 650 · Foreign-exchange practices, 651 · The rate of exchange, 654 · Specie parity, 654 · Purchasing-power parity, 655 · Speculative rates, 658 · Pegging the rate of exchange, 660.

XXXV. THE BALANCE OF TRADE AND INTERNATIONAL DEBTS 661

Our balance of trade before 1873, 661 · Our balance of trade as debtor country from 1873 to 1914, 663 · Increased exports during the World War, 664 · The United States as a creditor nation, 665 · Economic effects of an unfavorable balance, so called, 667 · Factors governing the balance of trade, 670 · The effect of reparations and war loans on international trade, 671 · War-debt settlements, 674 · The Dawes plan and the German reparation payments, 676.

XXXVI. THE LAW OF COMPARATIVE ADVANTAGE 681

The basic reason for international trade, 681 · Comparative cost in regional trade, 682 · Comparative cost in international trade, 686 · The menace of foreign competition, 694.

XXXVII. TARIFF AND FREE-TRADE PRINCIPLES 700

Tariff history, 700 · Arguments in support of the protective tariff, 703 · The infant-industry argument, 703 · The home-market argument, 704 · The high-wages argument, 705 · The national-defense argument, 706 · The validity of the arguments for protection examined, 706 · Industries injured by a tariff policy, 712 · British and American tariff policies, 714.

INDEX 719

PRINCIPLES OF ECONOMICS

BOOK I. PRODUCTION

CHAPTER I

INTRODUCTION

The Subject Matter of Economics. The study of economics is assuming a place of increasing importance. The disturbed conditions following the World War have enormously enhanced the interest of the general public in economic problems.

A similar situation existed a hundred years ago, after the Napoleonic wars; and the problems thrown up by that great struggle were in large part similar to those we face today. Sharp controversies arose over the causes of the conditions that prevailed and the proper remedies to be applied. British agriculture floundered in a tremendous depression, much like the one through which American agriculture has recently been passing. The landlord class demanded protection to shield it from the competition of the fertile land in the New World. British industry, on the other hand, wanted free trade, that it might buy raw materials in the cheapest markets. A struggle arose between these two contending groups. Each had its advocates; each sought to prove that Great Britain would benefit by its own special program. Two great economists in particular stand out in this controversy: Thomas Robert Malthus, who defended the position of the landlord class, and David Ricardo, who pointedly set forth arguments favorable to the position of the industrialists. The struggle became a leading political issue and finally ended in a victory for free trade and the manufacturing class.

Another controversy that emerged in the post-Napoleonic period was the currency question. England had temporarily abandoned the gold standard, and her paper money had depreciated. As a result there had been an enormous rise in prices. After the war these prices fell rapidly. Certain classes benefited from this; others suffered severely. What was the cause of the fluctuating value of money, and what could be done to stabilize the currency? What banking reforms were needed in order to prevent these violent fluctuations in prices?

During the war England had borrowed huge sums of money, and the public debt had reached unheard-of proportions. How were the funds for meeting the interest charges to be raised? How rapidly might the debt be paid off? Who ought to bear the brunt of the taxes? These were typical economic questions dominant in British politics of the time.

We are today faced with similar problems and many more besides, some of which are an aftermath of the World War; others are not. What were the causes of the post-war agricultural depression? How is the situation to be remedied? How can business be stabilized so that we may in the future escape such a train of bankruptcies as we suffered during the years 1921–1922? How shall we prevent a recurrence of such a period of unemployment as we witnessed in 1921–1922, when five million wage-earners could not find jobs? Would it have been better if the war had been financed more largely from taxes and less from loans? Should the excess-profits taxes have been continued? What are the effects on business prosperity of the high tax rates on large incomes? Do all forms of taxes tend to raise prices? Are freight rates too high? Should coöperative movements among the farmers be encouraged? Are trade associations good for the country? Are trusts and monopolies acquiring a strangle hold on business, or is competition still dominant in most fields? Should the United States Steel Corporation be dissolved? What caused the great increase in strikes in 1919? What is the solution of the labor problem? Why are wages so much higher than before the war? Ought we to pass the child-labor amendment to the Constitution? Are minimum-wage laws a good thing? Should trade unions be rooted out? Why are wages higher in the United States than in England? Should we restrict immigration into the United States? Should we be better off if we canceled the war debts? Why did the rates of foreign exchange fluctuate so violently in the post-war years? Is it desirable for a country to export more than it imports? Should people be urged to patronize home industry? Are the Federal Reserve banks a benefit to the country? Is our national income rising? Do retailers receive excessive profits? Should speculation in the stock market be prohibited? What determines the value of stocks and bonds? What determines prices, wage rates, interest rates, profits, and rents?

These and other questions are being discussed on all sides — in the newspapers and magazines, in hundreds and thousands of books, on the public platform, in legislative debates, and in private conversation. Every person in modern times is perforce more or less concerned with economics.

It must not be thought that any careful student of economics can give answers to all these questions. There is a consensus of opinion among economists on a great many questions, but many others are unsettled and there is no agreement concerning them. But even the unsettled problems will surely be approached more intelligently by those who have a firm grasp of economic principles. There are, moreover, many popular economic fallacies which can readily be exposed by any competent economist.

There are two methods of finding solutions for practical problems. Those who are impatient and "practical" will usually seek immediate results. They will study the most obvious phases of the problem and propose measures that may bring temporary relief. If their remedies do not work, they quickly turn to a different cure; and not infrequently the evils complained of are either forgotten or have disappeared before any settled conclusions emerge from such superficial experimentation. A scientific approach to economic problems demands that before any person proposes remedies he shall have made a thorough study of the fundamentals of economics. He must know how prices are made — what forces control the demand for and the supply of goods and services, and how these forces act. He must understand the operation of the broad social forces that control the division of the nation's income among landlords, capitalists, entrepreneurs, and wage-earners. He must have in mind a general idea of the way the whole machine works before he begins to tamper with a particular part of it.

Economics deals, as may be inferred from the enumeration of problems listed above, with a certain set of related phenomena, a certain aspect of human behavior. It is concerned with problems relating to the value and price of goods and services; with economic institutions such as property and contract, trade unions and corporations; and with the social control of economic behavior. All these matters are closely interrelated. They are not separate divisions in economics. They will not be treated as such in this book. Each is affected by the others. Value and price cannot be discussed and explained without due weight being given to economic institutions. It does not require a technical study of economics to see that trusts, combinations, trade associations, coöperative societies, trade unions, the speculative exchanges, and so on affect prices, profits, interest rates, wages, and rents. Similarly prices are affected by governmental policies — tariffs, "trust-busting," taxation, and regulation. Some economists indeed contend that price is the central problem of economics, that all the subject matter of economics is finally focused on the price problem.

If, however, we turn our attention to economic institutions or social control, we find a similar interrelation. Economic institutions are, in large part, the result of the contest between groups over price: each group striving to get a larger share of the national income, each trying to get a more favorable price for its product. Moreover, no society can consciously guide its economic processes in a wise manner without a knowledge of fundamental price and value principles. Finally, the various existing institutions that we have enumerated above affect directly or indirectly the problems of social control, and, on the other hand, social control eventually results in a modification of the prevailing economic institutions.

It is not necessary to trace these interrelations further. We have in a sense a price economics, an institutional economics, and an economics of social control. But there are no watertight compartments, nor can any one type of economic analysis claim primacy over the rest. All are tools to a common end: an understanding of economic life, of human behavior in its economic relations. And certainly the primary purpose of this knowledge is to point the way to the largest measure of economic welfare. All science is ultimately pragmatic. Economics is no exception. Adam Smith, the founder of modern political economy, whose great book was written in 1776, contended that "the object of the political economy of every nation is to increase the riches and power of that country." Dr. Pierson defines economics as "the science which teaches us what rules mankind should observe in order to advance in material prosperity." But knowledge must come first. Adam Smith emphasized this fact by the very title of his book, *An Inquiry into the Nature and Causes of the Wealth of Nations.* It is the function of any science to analyze, to find relationships, to explain; but the explanation is generally sought as a guide to action and in any event will often prove of unforeseen practical value.

So far as economic welfare is concerned, there are two primary problems confronting any society. The first has to do with the relative importance of different human wants; the second has to do with the most economical way of satisfying those wants. The first runs off into the field of ethics, of human values; the second is the primary concern of economics.

Economics as a Science. Economics assumes, as does all science, that nothing happens by chance, that any observed phenomenon of human life is contingent upon other phenomena, that with sufficient knowledge everything can be traced back to antecedent conditions. Science seeks to unravel the sequence of events. Of what does an explanation consist? It consists of pointing out what went before, of showing the sequential relations between phenomena. Certain results are observed. What causes brought them about? If the antecedent events are not fully known, a theory, or hypothesis, may be invented to explain the known effect. Such a theory, or hypothesis, consists of a statement of an assumed sequence of events which may or may not prove by later observation to be true. It is the aim of science to find out how and to what degree phenomena are related to each other; to what extent they are contingent upon each other. We may put it in a slightly different way and say that it is the function of a science to ascertain facts and to correlate these facts.

Now a large part of the phenomena of economics consists of human behavior in economic matters. Observation shows that on the average we can expect a mass of human beings in the aggregate to act pretty generally in a certain way under certain conditions. There will be, of course, a certain amount of variation from the normal. But repeated observations will show that even this variation can be predicted with fair accuracy. In all industrial nations a considerable proportion of the working classes form trade unions. A considerable percentage go on strike every year. The number of strikes varies with the conditions of business; but if we make due allowance for this, we can be fairly certain that approximately so many strikes will occur in any given country in any given year. On the average about 1 per cent of business firms go bankrupt in the United States every year, again varying with the fluctuations of good and bad times. Until immigration into the United States was restricted by law, the volume of immigration rose and fell with prosperity and depression. Approximately 70 per cent of the total national income in the United States goes to wages and salaries, and about 30 per cent goes to income from property. The percentage of the total national production which we export is approximately the same from year to year. If the price of cotton rises unduly compared to other things, the consumption of this commodity will be reduced. These illustrations could be multiplied indefinitely. In short, human beings *in the aggregate* under given conditions can be expected to act *on the average* in a certain way. But there will be considerable variation. The result is only an approximation, a probability.

There is no typical "economic man" who can with unvarying regularity be motivated by any cut-and-dried bundle of economic stimuli as a puppet is manipulated by a few strings. But we can be certain that, in the *aggregate* and on the *average*, human beings will act in a certain way when subjected to certain economic conditions. But there are other forces at work besides the economic ones, and the economic factors themselves are so numerous (and, moreover, frequently unknown and unmeasurable) that any conclusions as to results can only be true on the average and with a considerable though often measurable degree of variation. This variation does not indicate that things happen by chance, but rather that there is a multiplicity of causes at work.

Hence the conclusions of economics are often referred to as *tendencies* rather than as *laws*. The latter term, though used in economics, is on the whole more applicable to the natural sciences. Thus we speak of Boyle's law, the law of gravitation, and so on. These are apparently unvarying laws; within known limits the results are definitely predictable; the known and measurable conditions always lead to certain definite results. The term "tendency," on the other hand, implies that the resulting effect of certain known conditions is subject to considerable variation. Thus, with regard to large aggregates or masses of persons or events, certain results may be expected on the average with a certain measure of variation, but these results cannot be predicted so far as individual cases are concerned. The conclusion stated is true of the group but not of the individual. In the individual case the

Scie.

cause in question may be totally inoperative. A certain economic stimulus may completely fail to motivate certain individuals, but in the mass of human beings it may be expected to produce a certain average result.

The word "tendency" implies that there are so many different forces at work, some of which are not known or measurable, that the result attributable to a certain cause is obscured; the net resultant of the multiplicity of causes may be quite different from what it would be if the single cause upon which attention is centered were alone operative. The multiplicity of causes may produce results quite at variance with the one expected. Thus it sometimes appears that certain economic laws stoutly defended by all competent economists do not work out in practice. Consider, for example, the law of diminishing returns. As originally set forth, this law states that if certain natural resources and knowledge of the arts are given, every increase in population will tend to result in a larger *total* product but a constantly diminishing *per capita* product. The per capita production enormously increased, however, in the nineteenth century in spite of an unprecedented increase in population. The alleged result did not appear. Many persons have therefore concluded that we may safely discard the so-called law of diminishing returns and relegate it to the limbo of outworn philosophies. Closer examination will disclose, however, that this is a hasty conclusion. The law assumes a definitely limited supply of natural resources and no change in the arts of production. But the supply of available land was rapidly increasing in the nineteenth century, and the processes of production were rapidly being improved. The conditions being different from those stated in the law, the results would also obviously be different. Even though we had now reached the end of free land in the world, it might still prove true that the per capita product would continue to increase with increasing population because of inventions and improvements. Shall we then say that the law does not hold, or is at all events of no significance? Not at all. In physics we learn that if a certain force is applied to a body of a certain weight and mass in a certain direction, the body will move in that direction with a certain definite velocity. A man is rowing a boat on a river, yet with reference to the bank the boat is not moving at all. We discover that the man is rowing against a strong current. One force is counteracted by another force operating in the opposite direction. Obviously it would be quite incorrect to say that the law referred to above is not operating. Instead of a simple set of conditions we have a complex set of forces. A definite result must inevitably follow, but the result is different from what it would be if only one set of conditions obtained. Yet we cannot understand the result without analyzing the effect of each factor separately. Because of the complexity of forces at work it is probably safe to say that no law of the physical universe ever appears to work out as stated in scientific laws. If the result observed in the actual world always corresponded with exact precision to the result predicted by science, it would probably not be necessary to engage in painstaking research to discover scientific laws. It is because of the complexity of factors at work that all scientific laws when first promulgated appear startling and unreal. Bodies do not actually fall according to the law of gravitation, for they are retarded by atmospheric pressure. The result pointed to does not seem to occur in actual life. It is no wonder that so-called practical men have little faith in theory.

Thus in a sense it may be said that all scientific laws are tendencies. But there is this difference: the conclusions of scientific study may properly be called laws if each set of conditions is subject to control so that each separate force can be measured and the resulting effect definitely predicted. In economics, on the other hand, many of the forces at work are often altogether unknown, others cannot be measured, and therefore the resulting effect of those forces which can be measured is subject to considerable variation, since it is vitiated by the unknown or the unmeasurable forces. Thus *exact* prediction becomes impossible, though prediction within certain limits of variation is often possible. In such cases the term "tendency" is to be preferred to the term "law."

An illustration may help to clarify this matter. The accompanying chart (Fig. 1) shows the relationship between the yield of corn and the price per bushel (the trends having been eliminated in both cases). The curve represents the average relation between these two variables. If the yield is low, the price tends to be high; and if the yield is large, the price tends to be low. But the curve represents only an average condition, an approximation, a probability, or tendency. It is impossible to say that with a certain yield the price will be exactly so much, because there are many other factors besides yield which affect price, and which have not been adequately taken account of. The curve tells you what approximately may be expected in the aggregate, but the individual instance may be quite different. Moreover, the chart also tells you the degree of variation that may be expected. Though the individual price may not conform to what may on the average be expected, we can be reasonably certain at least that it will not vary more than so much. That is a matter of no small

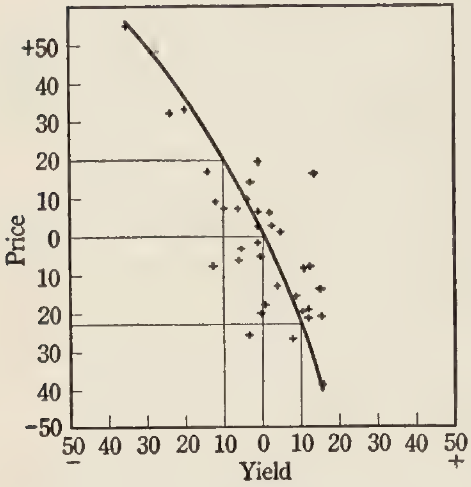

FIG. 1. The relationship between the yield of corn and the price per bushel

The scattered points indicate the *actual* relationship between yield and price. The curve is fitted to these points and therefore represents the *average* relationship between these two variables. When the yield is low the price is high, and vice versa. The zero point, on the yield scale, represents a normal yield, and, on the price scale, a normal price. The figures on either side of zero represent deviations above (+) or below (−) normal. From the diagram it will be noted that if the yield is 10 per cent above normal, the price is about 20 per cent below normal; if the yield is 10 per cent below normal, the price is about 20 per cent above normal. (See Moore's *Generating Economic Cycles*, p. 36) importance. In so far as economics can measure the probable result of certain conditions, together with the approximate variation from this result, it may properly be said to be a true science. In the main, economic laws are of this statistical type. They hold for aggregates but not for individual instances. They set forth the average relationship between phenomena. They are approximations or tendencies.

We often hear people talk about the "natural" laws of economics, especially the law of supply and demand. We cannot alter, so we are told, the law of supply and demand. Now this statement indicates both a great truth and a great misconception. It contains a truth because it implies that a scientific law or tendency cannot be counteracted by blindly ignoring it, that the results follow from definite causes. It contains a great misconception, however, because it gives one the impression that human beings are helpless in the face of scientific laws. If certain conditions obtain, human beings are indeed helpless to avert results that follow with a certain measurable variation therefrom. But the conditions themselves may be changed; and when they are, other results of a different sort will then inevitably follow. In this respect economics and other social sciences do not differ from the natural sciences. Because certain conditions inevitably result in certain diseases, we need not submit to disease in despair. We can alter the conditions and thus change the results. Economic conditions are especially subject to alteration because they are in part man-made institutions. Our economic life is therefore subject to human control; it is not at all a blind play of uncontrollable external forces.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "A scientific law states that definite causes necessarily lead to definite results." — SELIGMAN. Comment.

2. "Everything that occurs in economic life takes place in accordance with some law; it is the function of the economist to ascertain that law." — SELIGMAN. Criticize.

3. "An economic law is a natural law so far as it states that given conditions will lead to given results. An economic law is not a natural law so far as it implies that human effort is impotent to modify the conditions which lead to the results." — SELIGMAN. Do you agree?

4. (1) The purpose of economics is to *explain* the phenomena of economic life. (2) The purpose of economics is to discover ways of attaining a larger measure of economic welfare. With which of these statements do you agree?

5. "We cannot in practice consider a fact without relating it to other facts, and the relation is a theory." — CLAY. Comment.

6. "The man who opens a discussion by saying that he is going to 'deal with facts not theories' does not mean that he is going to refrain from generalization; he usually means that he wishes to confine attention to a few facts that support his generalization, and to ignore all facts that conflict with it." — CLAY. Is this true?

7. Students often remark, "That may be true in theory, but not in practice." Is this possible? What is probably really meant by the statement?

8. What do you think of the following remark by a cynical critic of social reform. "Many people foolishly think they can substitute the laws of Congress for the laws of economics."

9. "He [Herbert Spencer] taught me to look on all social institutions exactly as if they were plants or animals — things that could be observed, classified and explained, and the action of which could to some extent be foretold if one knew enough about them." — BEATRICE POTTER WEBB. Is this view sound?

CHAPTER II

PRODUCTION

Wants and Production. Human beings have wants. Both as individuals and as members of social groups they are constantly striving to satisfy these wants. Wants are satisfied in part by the manipulation and utilization of the forces and properties of nature, in part by the direct rendering of personal services. "Production," as the word is used in economics, means any human activity that results in the satisfaction of human wants. In the modern exchange economy people are willing, either individually or collectively, to pay a price for such acts, but of course no one would be willing to pay a price for them did he not *anticipate* that they would result in the satisfaction of human wants. Such activities are *productive*, and persons engaged in them are said to be *producers.*

These activities may give rise either to concrete tangible goods, such as the merchandise displayed in shops, or in direct personal benefits, such as the transportation of people by railroads, the work of physicians, entertainers, and government officials. The producers of these goods and services are primarily concerned with getting a price for their activities, but the result of their efforts is nevertheless, in the usual and normal case, the satisfaction of human wants.

Some productive activities may not be desirable from the social point of view and in the opinion of moralists; but so long as someone is willing and able to pay for them, they are *economically* productive. Instances in which the moral and the economic criteria come into conflict are numerous. The manufacturer and the seller of adulterated goods or of patent-medicine nostrums are, in an economic sense, productive so long as they can find buyers. That their activities are in an overwhelming majority of cases socially pernicious most persons would at once concede. Hence laws are enacted to restrain these types of production. But since a price can be obtained for such wares, the makers and venders of them are engaged in production. The economist's position in this matter is not based upon a belief that economics has nothing to do with ethics. The effect of ethical standards upon all the affairs of life is too great to permit the consistent maintenance of such a belief. But if once one concedes that unethical acts are not productive, he has entered upon a course of reasoning which will lead to the conclusion that a very large part of human endeavor is unproductive from the point of view of different groups of people. The illustrations that we have given are cases in which a majority would probably agree that production and sale were socially harmful. But what about the writing and sale of cheap, salacious fiction? or the production and exhibition of motion pictures that depict acts of violence or crimes? or the sale of whisky in countries where no legal prohibition exists? Certain groups of people in India have religious scruples against the use of beef for human food, but most of the world has not. Now economics strives to set up principles that shall be, as nearly as possible, generally applicable to all economic relations. The scope of its investigations cannot, therefore, be limited to occupations and acts that receive unanimous ethical approval. Moreover, since it deals primarily with the valuation of goods, it cannot refuse to consider valuations which concern commodities that harm rather than benefit men.

Economics is not, however, estopped from considering the effects of certain types of production or of consumption upon social welfare. If the consumption of alcoholic beverages affects the production of wealth adversely, that is an economic fact and cannot be ignored. Or if industrial education in the public schools leads to greater efficiency in production, that fact must be given its proper weight in considering the reasons why a nation that provides such education can undersell another nation that does not. But economics must consider these conditions only in so far as they affect production, valuation, and distribution, and not as moral or educational problems.

Not all wants are satisfied through the activity of human beings; many are satisfied without any effort at all. The air we breathe satisfies an absolutely essential need, without a thought on our part. But a large range of human wants cannot be satisfied without effort. Even when we depend upon the forces and properties of nature, in large part these forces and properties must be adapted to our needs before our wants can be satisfied. Because most of our wants would otherwise go unsatisfied, we are willing to engage in productive effort.

Classification of Productive Activities. Productive activities may be classified under the following heads:

1. Agriculture
2. Lumbering
3. Mining
4. Manufacturing
5. Transportation
6. Exchange
   - *a*. Merchandising
   - *b*. Banking
7. Personal and professional services.

The term "production" was formerly not used to include all the activities listed above. Even today the term is used popularly in a much narrower sense. Persons engaged in agriculture, lumbering, mining, and manufacturing are often spoken of as if they were the only "producers." When men produced exclusively for their family wants, or even later when they produced for a purely local market, there were indeed few other forms of productive activity. But with the widening of the market and with the coming of territorial specialization, when the people of one region devoted themselves to the production of one commodity and those of other regions to other commodities, a considerable amount of labor was of necessity employed in transportation and exchange. As the market became wider and the geographical specialization of labor increased, the producers of raw materials and manufactured products became increasingly separated from the consumers. This separation had somehow to be bridged by transportation, merchandising, and banking activities. Without these activities we should be forced back into the primitive economy of small local markets. Such an economy would mean that we should lose all the gains of territorial specialization, since every region would be compelled to produce for itself all it needed. It would also mean that we should have to forgo all the advantages of large-scale production, since a small local market could not possibly absorb all the output of a large-scale plant. In place of the modern machine method of production we should have hand tools, small shops, and small-scale production. All this would entail an enormous loss of want-satisfying goods. And yet this is exactly where we should be were it not for the functions carried on by transportation and exchange. It must therefore be evident that labor engaged in transportation and exchange is a necessary part of a whole group of activities which results in yielding satisfactions and utilities. Such labor is therefore just as truly productive as labor involved in agriculture, mining, lumbering, or manufacturing. Any activity, other than play, which in any way contributes to the satisfaction of human wants is productive.

It is certainly true that satisfaction-rendering objects (goods) may be produced without the aid of middlemen, merchants, bankers, or even transportation workers. They were so produced in large part during the handicraft age. But they could not be produced *in the manner* in which they are produced in modern times without the aid of these groups, and therefore not so cheaply. Transportation and exchange are vital and necessary parts of the modern method of production. The modern method of production is a complex chain of activities in which there are many links. Goods are no longer produced by one person or even by a special group of producers. Goods are produced by a society practicing extreme geographical specialization and division of labor. Any firm or persons constituting a link in that complex and specialized process are certainly engaged in the production of goods.

In the handicraft period the market was, as we have already noted, largely local. The exchange of products between crafts in the local market was a simple matter. The master craftsman was therefore worker and merchant, all in one. In the present economic order, with national and world-wide markets, the exchange of products between different regions, different industries, and different stages in the process of manufacture becomes an exceedingly complex matter. Each stage in the process is dependent on the firms in the preceding stage for the supply of raw material, and on the next stage for the market for its finished product. Each group of producers is dependent upon some other group of producers for a market for its products, and each group of producers in turn constitutes the consumers of products manufactured by other groups of producers. These various groups of producers and consumers must be brought in contact with one another through the processes of transportation, exchange, wholesaling, and retailing.[1]

It thus becomes evident that the country elevator man, the warehouseman, the wholesaler, the retailer, and the banker are just as truly engaged in production as are the farmer, the miller, and the baker. Some economists contend that the baker who transforms flour and other raw materials into bread creates *form utility*; the teamster who hauls the bread to the retailer creates *place utility*; the elevator man who stores the wheat creates *time utility*; and the retailer who sells the bread creates *possession utility*, since he transfers it to persons to whom the possession of bread is of greater utility than to those from whom he purchased it. But this is merely a way of stating that each specialized worker plays a part in the elaborate process by which our modern urban populations are furnished with bread. It is very artificial to say that one produces form utility, another time utility, and so on. Each coöperates in a complicated process which finally results in placing bread in the hands of consumers, and that is all there is to it. Thus in the production of bread we have the following coöperating specialists: farmer, elevator man, railroad crews, manufacturer of railroad cars, coal and iron miners, iron and steel workers, lumbermen, wholesale merchant, miller, banker, truckman, baker, retailer, and many others.[1]

---

[1] For a fuller discussion of the marketing function in production see Chapter VI of this book.

Early writers on economics also thought of persons rendering personal and professional services as unproductive laborers. A little reflection will, however, soon convince one that it is very illogical to call the maker of a surgical instrument a producer and refuse to admit that the surgeon who uses the instrument is a producer. A violin in the hands of a great artist renders satisfaction to wants. The flow of utilities thus

---

[1] The following tables give some illustrations of the relative importance of these various functions in the modern economic order. (See Report of the Joint Commission of Agricultural Inquiry, Part IV, 1922.)

1. Distribution of the Dollar paid by Consumers for Bread

	1913	1921
Farmer	28.0	28.1
Transportation	2.8	2.6
Elevator	1.1	2.8
Flour manufacturer	0.6	0.6
Transportation	2.4	4.4
Bread manufacturer	43.6	42.9*
Retailer	21.5	18.6
	100.0	100.0

* Includes selling cost.

2. Distribution of the Dollar paid by Consumers for Men's Shoes

	1913	1921
Raw materials	40.8	38.5
Manufacturer	30.5	33.2
Retailer	28.7	28.3
	100.0	100.0

3. Growers' and Distributors' Returns on California Oranges per Box

	Five-year average 1913–1918	December 1, 1920, to November 15, 1921
Growers received for fruit on tree	$2.27	$2.30
Harvesting, packing, and selling	0.51	0.88
Transportation	0.94	1.62
Wholesalers	0.48	0.61
Retailer	1.50	2.06
Consumer paid	$5.70	$7.47

rendered is produced both by the maker of the violin and by the artist who plays it. Both are producers, for their activities directly or indirectly render satisfactions, utilities, services. All activity that results in the satisfaction of wants, and for which someone will pay a price, is productive.

The Source of Utilities. The satisfactions, utilities, or services that come to us have various sources. Many of our wants are satisfied by nature directly, without any activity of man at all. Illustrations at once suggest themselves: the air we breathe, the heat and light from the sun, the running brook, the fruit on the trees, the beauties of rocks and mountains, of valleys and the setting sun. Other wants are satisfied by the action of man engaged in fashioning, shaping, or transforming the objects of nature in such a way that they will be capable of yielding satisfactions. Thus materials are taken from mine and forest and fashioned into houses, furniture, and musical instruments. In his labor man is constantly utilizing forces of nature which assist him in his work. The forces of waterfall and wind, heat and electricity, are harnessed to assist him in his work. Other wants are satisfied directly by the activity of human beings. As examples we may cite the utilities rendered by servants, barbers, preachers, teachers, singers, and actors.

Economic Utilities. Some of the utilities referred to above are rendered by nature in such abundance that the wants to which they minister are fully satisfied. For example, air exists in such quantity that all want for it is completely satisfied. But most of the want-satisfying services rendered by nature and man are scarce. They are not offered in sufficient abundance to satisfy all want. If one unit of these services were lost, a certain amount of human satisfaction would be lost. Each unit matters; it has significance, importance, worth. Because of this we act economically toward those services which are scarce. We act on the assumption that there is not enough of them to go around. We economize. We therefore call them *economic* utilities, services, or satisfactions.

But we would not act economically toward them if we were able in no way to *control* the utilities, services, and satisfactions rendered. In the northern climate the heat from the rays of the sun is not sufficient to satisfy human wants, but the utilities derived from them are nevertheless not economic utilities because nobody has so far been able to control in any way the rendering of these utilities. Since we cannot control the rendering of these utilities, it is impossible for us to act economically toward them; and they are therefore not economic utilities. We are concerned in our study particularly with those utilities, satisfactions, and services which are *scarce*, which we therefore wish to act economically toward, and which we *can act economically toward because we are able to control them.* The importance of this question of control will become apparent later when we consider the subject of property and income.

It has sometimes been asserted that the test as to whether utilities, satisfactions, or services are economic or not depends upon whether or not human labor has been applied to them. It must be obvious from the foregoing discussion that this is not the case. Many utilities derived directly from nature without any assistance on the part of man are economic. The utility afforded by the surface of the earth as a place to stand upon and to build upon is an example; likewise the utilities rendered by lakes and rivers, shade trees (in some parts), wild fruit trees, and wild animals. Economic utilities are not necessarily created either in whole or in part by labor. They are derived in part from nature, in part from natural objects fashioned and transformed by man, and in part from man directly.

Production in the Modern Economic Order. It is now clear that the flow of utilities, services, and satisfactions which constitutes our real national income comes in part from nature and in part from the productive services of man. We shall now attempt to analyze the way in which productive services are rendered in the modern economic order. To make this analysis clear let us briefly contrast the modern method of production with the method of the handicraft period. In the handicraft period there was little specialization or division of labor within the industry. The master craftsman or the journeyman performed all the various processes involved. There were no complex and highly specialized agents to be coördinated, as is true under the present system. Contrast this method of production with the modern process of making shoes, for example. Here we have a great variety of highly specialized machines, and every step in the process requires a different kind of labor. These minutely specialized processes must be properly coördinated. Foremen, managers, and superintendents are needed. But this is not sufficient. Not only must we have many types of manual laborers, clerks, technical experts, accountants, managers, and superintendents, but also some person or group of persons must promote the business in the first place and later assume the ultimate control and responsibility for the successful carrying-on of the venture. There must be "captains of industry," or *entrepreneurs*.[1]

In the modern economy it is not always easy to locate the *persons* who perform the function of entrepreneurship. In small-scale businesses the individual proprietors are of course the entrepreneurs; but in large-scale businesses the function is performed by a business organization, such as a corporation, or, rather, it is performed by persons functioning through such an organization.

It must not be assumed, however, from what is said above that in the primitive local economy the entrepreneurship function was lacking. But it is certainly true that in the modern economy, with its complex system of large-scale production, the function of entrepreneurship is far more important than was the case in the days of local self-sufficiency. The modern entrepreneur stands at the center of the economic organization. All the specialized agents of production are under his control. It is he who sets the productive resources in motion.

---

[1] For a fuller discussion of the function of entrepreneurship see Chapter IV of this book.

In the next place it may be observed that in the handicraft period production was carried on by a short-time direct process. The tools used were suitable devices to assist the work of the hands. Little work was involved in making these tools compared with the work employed in making the finished product. In the modern system production is carried on by a roundabout, time-consuming process. Much more labor is involved in the production of machines, factories, railroads, and so on than is involved in the direct production of the consumption goods that at last satisfy human wants. When machinery has been built, it is offered for sale by the manufacturing company. From the standpoint of economic theory we may conceive of this machinery as a bundle of uses which may be obtained from it through a succession of years. These uses will assist in the production of consumption goods which will finally satisfy human wants. It is because the machinery will indirectly but ultimately aid in satisfying human wants that it has value. It may therefore be thought of as a bundle of future utilities. But no one prizes future utilities as highly as he does present utilities. Therefore the present value of the machinery must be considerably less than the sum of the values of the utilities it will ultimately render. Now if the manufacturing company is to be able to sell this machinery, someone must be willing to give up present purchasing power in exchange for this machinery, which is in reality a bundle of future utilities. Indeed the machinery would never have been built, had not the manufacturing company been fairly certain that someone would be willing to give up present purchasing power for this bundle of future utilities. But no one would do so were it not that the sum of the bundle of future utilities which will finally be received will probably be more than the utilities represented by the purchasing power given up. The person who gives up present purchasing power in exchange for the machinery is to that extent forgoing immediate satisfaction of wants. Since capital equipment would never be produced unless someone was willing and able to buy it and hold it until its future utilities were forthcoming, and since no one can buy it without giving up present satisfactions, it is obvious that under the modern system of production, carried on as it is by means of elaborate capital equipment, this service of waiting or forgoing present goods, which takes the form of saving and investing, is an absolute necessity.

The services rendered under the modern system of production may then be summarized as follows:

1. The services of labor
    a. Manual labor
    b. Mental and technical labor
2. The services of entrepreneurship
3. The services of waiting, saving, investing
4. The services of nature (natural resources)

Control over Utilities. We have so far considered the source of the continuous flow of utilities, satisfactions, or services. We shall now consider the means whereby in modern society control is exercised over this flow of utilities and services. In so far as the utilities are subject to that sort of control which can be bought and sold, which can be exchanged in the market, we have the emergence of property. Property, the court says, is anything which has exchange value, and it would not have value if it did not give control over utilities. Property is therefore the legal claim which gives transferable control over utilities and services. In so far as utilities are not subject to that sort of control which is transferable, the control is exercised through the institution of personal freedom. Since slavery was abolished, the law has not permitted one person to transfer to another permanent control over the utilities his person is capable of rendering.

A portion of personal and professional services is controlled directly by the persons rendering them. The services rendered by the servant form the clearest illustration. But the proportion of utilities thus received by consumers is very small. Take the case, for example, of the motion-picture theater. The utilities rendered in this instance are a product of a whole network of contributing factors. There is the theater building with all its furnishings, and the motion-picture machine with all the labor, saving, entrepreneurship, and natural factors employed in its production. There is the film with all the actors and other persons who helped to produce it. Whole industries assist in the process of furnishing motion pictures to the public. The "movie" actors do not sell motion-picture utilities to the public. The utilities rendered are under the control of the entrepreneur, or rather a group of entrepreneurs.

Control implies the power of *withholding* these utilities, services, and satisfactions. The power to withhold gives the bargaining power to demand services in return for services rendered. The flow of utilities, satisfactions, and services derived from this control constitutes one's income. This exchange of services against services takes place on the basis of the importance of each. The significance of each service, utility, or satisfaction is measured in terms of other utilities, services, or satisfactions. This is the process of valuation. Thus the value of property arises from the flow of utilities and services which may be obtained in exchange for the services over which it gives control.

Now our motion-picture entrepreneurs have the power to yield or withhold the utilities of motion pictures. They are therefore able to command a flow of utilities in exchange for them or, what amounts to the same thing, to command a claim (in the form of money income) upon a flow of utilities. True, they are compelled to share this claim upon a flow of utilities with all the contributing factors employed by them: the laborers, the savers, the owners of natural resources. If competition has been at work long enough to have attained its full influence, the entrepreneurs will be compelled to share all the money income with the contributing factors, reserving only enough for themselves to compensate for their own contribution. But competition may not be at all effective. The motion-picture entrepreneurs may be sheltered from competition by patents, secret processes, or combination, or goodwill, so that they will not be compelled to hand over all the money income to the contributing factors. Such control gives rise to a kind of property which we designate "intangible property."

The foregoing discussion may be summarized as follows:

1. Property is the institution which gives transferable control over utilities and services.
2. Personal freedom is the institution which gives nontransferable control over the utilities and services which may be rendered by one's own person.
3. Income is the flow of utilities or services which one can command because of control over the rendering of other utilities and services.
4. Money income represents a claim upon the flow of utilities or services which one can command because of control over other utilities and services.
5. Intangible property represents transferable control over utilities or services rendered under conditions which permit of *unusual* withholding power, such as may be derived from patents, secret processes, goodwill, or combinations. The excessive power thus derived to withhold the utilities or services means excessive power to demand services in exchange for services rendered. This excessive power is represented by "intangible assets."
6. This outline indicates the interrelations mentioned above:

$$\left.\begin{array}{l}\text{Exchangeable control (property)}\\ \quad\text{Tangible property}\\ \qquad\text{Capital goods}\\ \qquad\text{Natural resources}\\ \quad\text{Intangible property}\\ \text{Non-exchangeable control (personal freedom)}\end{array}\right\} \left.\begin{array}{c}\text{Flow of utilities}\\ \text{and services}\end{array}\right\} \text{Money income}$$

Private Income versus Social Income. Patents, secret processes, and monopoly control are illustrations of intangible-property rights. They give the owner the power to limit output and thus secure a *larger price* for a product made scarcer by these restrictive devices. It may even be that by the use of these methods he is able to get a *larger income* from a smaller output of goods. Is he then more or less productive? From his own individual point of view the smaller production yields the larger income, but from the standpoint of society as a whole he is less productive because he produces a smaller quantity of want-satisfying goods. It is true that in the modern exchange economy no act is productive, from an economic standpoint, which does not command a price; but we cannot measure the *extent* to which an act is productive, from the social point of view, by the price that its products are able to fetch in the market. And why? Because while the product in question is enhanced in value by restriction of output, the value of other things is thereby lowered, the total volume of production is lessened, and the total real income of society is reduced. The final test of the *degree* to which an act is productive is the volume of economic goods and utilities resulting therefrom.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Material goods and personal services which render utilities that are *scarce* are referred to in economic literature as *economic goods.* How do you distinguish between economic goods and free goods?

2. Economic goods may be defined as goods which (1) yield utility, (2) are scarce, (3) are subject to human control. Defend this statement.

3. The term "property" is sometimes said to be synonymous with the term "economic goods." Criticize.

4. "Property" is sometimes defined as "material economic goods." Criticize.

5. The early writers on economics held the view that only such laborers as turned out material goods were productive. Adam Smith placed in the class of "unproductive" laborers "some both of the gravest and most important and some of the most frivolous professions; churchmen, lawyers, physicians, men of letters of all kinds; players, buffoons, musicians, opera singers, opera dancers." Is this a useful distinction?

6. In the terminology of the business world economic activities are divided into production, marketing, transportation, finance, and so on. In economic terminology the word "production" includes all these activities. There is something to be said in favor of both uses. Discuss.

7. "The view that production means the creation of satisfactions or utilities is naïve and old-fashioned; production in point of fact means doing anything for which people are willing to pay a price." Criticize.

8. "The real test as to whether labor is productive is this: Does it earn an income? If so, it is productive." Do you agree? Is the purveyor of quack medicines productive? the professional gambler? the seller of worthless mining stock?

9. Is there a sound distinction between the terms "criminal labor," "predatory labor," "acquisitive labor," and "productive labor"?

10. Productive labor, in the business sense, means labor which yields a money income; productive labor from the standpoint of human welfare means labor which has social utility. Which position should the economist take? Why?

11. A monopolistic combination restricts production and so raises the price. As a result it increases the total value of its products, but decreases the physical quantity of its products. Is it more or less productive? Explain.

12. "As has been intimated more than once in the course of this discussion, the limitation of supply to enhance the price is a normal and legitimate business practice by business men. . . . But modern trade-unionism partakes of the nature of business enterprise which implies that its members are more interested in wages — price per hour, per day, per week — than in output of goods — tons and yards per day, per week, per month." — ARDZOONI. If this is so, are trade unions productive?

CHAPTER III

THE REGULATION OF PRODUCTION

The Individualistic Order. Adam Smith, who wrote in 1776, just at the beginning of the Industrial Revolution, was an ardent advocate of laissez-faire. By that he meant an economic order in which there is a minimum of governmental regulation of industry and in which each individual is left free to engage in such business or calling as he may wish, to accumulate and use private property without undue restrictions by government, to enter freely into such contracts as he may desire without restrictions by law, to exchange products on the market at such prices as he can get. There are therefore four fundamental institutions which form the basis of a laissez-faire or individualistic economic order: (1) freedom of enterprise, (2) private property, (3) freedom of contract, and (4) freedom of exchange.

Such an economic order is sometimes spoken of as an unregulated society, but in fact this is not the case. The social control of economic life consists not only of deliberate governmental regulation, such as the regulation of railroad rates by the Interstate Commerce Commission, but also of such institutions as private property, freedom of enterprise, freedom of contract, and freedom of exchange. None of these institutions can exist without some form of social organization to give them effect, and all are products of the stage of social evolution in which we now live. All have important regulative effects upon our economic activity. A society, however, which is subject to no sort of regulation, if indeed such a society could exist, is completely anarchistic. The institutions of private property, free contract, freedom of exchange, and freedom of enterprise are maintained and enforced by the sovereign power of the state and are themselves great regulative forces in the economic world.

The Law of Supply and Demand. An individualistic order based on the four institutions named is often said to be regulated by "natural" or "automatic" forces, in contradistinction to the regulation imposed by conscious and deliberate social control. Thus in a society with private property, free contract, freedom of enterprise, and complete freedom of exchange there is no governmental body deliberately determining how many bushels of wheat or how many pairs of shoes shall be produced or the price at which these goods shall be sold. These matters are settled, it is said, by the natural and automatic working of the law of supply and demand. If there is a shortage in the production of wheat, the price will be exceptionally high and farmers will thereby be encouraged to produce more wheat. If plumbers are scarce, the wages of this class of workmen will be higher than for others of similar skill and consequently workmen will be attracted to this field. Under the automatic operation of the law of supply and demand market price becomes the guide, indicating how much of each thing ought to be produced and how the working population should be distributed over the field of industry.

This regulative principle is, however, only a long-run tendency. It assumes the swift mobility of labor and capital. But capital invested in a certain plant cannot readily flow into other more remunerative fields. Farmers cannot instantly change from one crop to another, nor can they easily transfer their capital from farming to manufacturing when they find that the prices of farm products are unduly low compared with the prices of manufactured goods, as has been true during the recent agricultural depression. Nor can wage-earners in other fields readily turn plasterers when plasterers' wages are phenomenally high, as was true in some cities a few years ago.

Assumptions of the Individualistic Philosophy. It is assumed by those who have complete faith in the automatic working of the law of supply and demand that market price can always be relied upon to indicate what ought to be produced. But as the French economist Gide has well said, the law of supply and demand distributes production not according to man's true needs but according to his desire and the price he is willing and able to pay. Now there are many moralists who are ready to assert that people buy ugly things when they ought to buy things which are artistic and beautiful; that they spend too much money on automobiles, motion pictures, confectionery, and cosmetics and not enough on education; that the rich through their purchases direct production to luxuries which have very little social utility, whereas the productive power of the nation might much more wisely be spent in supplying things that would raise the standard of living and the material well-being of the masses. It is thus claimed that the law of supply and demand is a very imperfect guide to production for two reasons: first, because income is unequally distributed and so luxuries are produced for the few even though the necessities and comforts of life are sadly lacking for the many; second, because all buyers, both rich and poor, have wants and desires that do not conform to the canons of the moralists. Thus it is asserted that a more equitable distribution of wealth and income would result in a more rational demand for goods and, furthermore, that education, culture, and morality would raise the standard of consumption.

The individualistic philosophy assumes, moreover, that everyone will follow his own self-interest in a rational way and that if everyone does follow his own self-interest, the greatest social good will result. But neither assumption is wholly tenable. People are too ignorant and easy-going to follow their true self-interest. Unless restrained by law they buy adulterated foods, worthless mining stock, etc. Workmen do not refuse to work in plants where the machinery is improperly safeguarded. People are induced by expensive advertising and high-pressure salesmen to buy things they do not really want. Nor does the self-interest of each lead to the good of all, particularly in a society of unequals. The strong and clever are all too likely to exploit the weak and ignorant.

The so-called natural or automatic working of the law of supply and demand implies, moreover, perfectly free competition. It assumes that there will be no artificial restriction of supply by means of monopolistic combinations. It assumes that if prices are unusually high in a certain field of industry, the large profits resulting therefrom will attract more capital and labor, production will increase, and the price will come down to a fair level in line with the prices of other commodities. Thus it is expected that competition and not governmental regulation will operate to insure a fair price. But in point of fact it is as natural to combine as it is to compete. And so we have our trade associations, monopolistic combinations, trade unions, farmers' organizations, professional associations, and the like. Self-interest impels competitors to combine.

Public Regulation. Modern societies have all reached the point where automatic regulation of the economic order breaks down. If the stage of combination is reached, the public in self-defense can no longer allow economic affairs to run their own course. Laissez-faire no longer gives adequate protection to the public interest. The public, through the state, must take hold and apply conscious and deliberate regulative measures. This may be done in one of two ways: first, measures may be adopted which aim to maintain competition and prevent combination; second, the combination may be accepted as inevitable, and measures may be adopted which seek to control the combination in the interest of the public welfare. One such measure is the regulation of prices. Such a measure aims to secure a "fair" price and adequate production. Prices may be regulated directly, as is done, for example, in the case of railroad rates; or they may be regulated indirectly by watchful oversight, as is done in the case of the packing and steel industries. Another measure by which regulation could be applied would be by taxation of the surplus incomes accruing from combination. While this method might insure a fair profit, it would not insure a fair price or a proper amount of production, because monopoly prices necessarily involve restriction of output.

Unregulated Private Capitalism versus Regulated Capitalism. Adam Smith and the early classical economists believed that the automatic regulative factors of private property, free contract, freedom of enterprise, and free exchange were sufficient to insure economic welfare. Such has not proved to be the case. No doubt these regulative principles would work out satisfactorily if human beings had reasonable equality of bargaining power and if they had the knowledge and intelligence to see the various alternative opportunities, and the economic independence to withhold their services and goods with a view to making use of the best available opportunity. But human beings are grossly unequal in knowledge and power, and the inevitable result was the oppression of the weak by the strong. Thus England found it necessary early in the nineteenth century to abandon the policy of relying exclusively on the automatic principles enumerated above, the policy of laissez-faire, the policy of a minimum of governmental interference in industrial affairs. Individual freedom of contract was restricted by governmental interference in the interest of the public welfare. In 1802 a law was passed prohibiting the employment of apprenticed children under nine years of age in factories and limiting the hours of labor to twelve for apprentices under sixteen years of age. This act applied chiefly to pauper children who had been apprenticed to mill owners by the parishes whose charges they were. With the coming of steam power the mills could be located in populous districts and it was no longer necessary to transport children (chiefly apprentices) to the remote places where water power was available. Hence apprenticed children were decreasing in importance and children living at home became of greater importance as factory labor. Accordingly, in 1819 the provisions of the earlier act were extended to all children working in the cotton mills, including those who lived at home and who were not apprenticed. Both these acts were left to local enforcement and were therefore largely ineffective. In 1833 an act was passed applying to all textile establishments except silk mills. This act completely forbade the employment of children under nine and limited the hours of labor for children under thirteen to nine a day and forty-eight a week. The act prescribed two hours of schooling per day for child laborers. For the first time reasonably effective enforcement was made possible by the appointment of factory inspectors. In 1842, following a Parliamentary investigation revealing shocking conditions in the employment of women and children in the coal mines, an act was passed which prohibited the employment of women and children underground. In 1847 the Ten Hours Law was passed, which restricted the hours of labor for women and children in textile mills to fifty-eight a week. These early acts themselves bear witness that the laissez-faire policy resulted in the exploitation of the weak by the strong and failed to secure the nation's health, morals, and general welfare.

The policy of social control has been carried much farther during the last seventy-five years. The act of 1864 contained provisions for sanitation. The act of 1867 extended the existing laws to include practically all kinds of factories and workshops. In 1901 the minimum age for child workers was raised to twelve years. In 1909 a minimum-wage law was passed by which wages in a few low-paid industries were fixed by Trade Boards. The scope of this act has been greatly extended, being made to apply to coal mines in 1912, and in 1918 extended to any industry where labor is not well organized and wages are relatively low. This act now covers about three million workers, largely women. Legislation was passed in 1897 by which employers are compelled to compensate workmen for industrial accidents. In 1911 compulsory insurance against sickness and unemployment was introduced and the employer was compelled to carry part of the expense burden. In 1909 the Housing and Town Planning Act was passed, which enabled local public authorities to tear down unsanitary tenements and to purchase land and erect clean and wholesome dwellings. Following the World War an extensive housing program was undertaken and unemployment insurance was extended to include practically all industries. This review of British social legislation is by no means complete, but is sufficient to illustrate the general tendency.

Similarly in the United States the automatic regulative forces operating by themselves alone have been found wanting. From the very beginning it was found necessary to regulate the banking business. In this matter free competition, freedom of enterprise, and self-interest were not sufficient. We have found it necessary to adopt a national program for the conservation of our natural resources. Freedom of enterprise and competition proved wasteful. We no longer permit immigration to be controlled by automatic forces, but definitely restrict the numbers that may come in from Europe, Australia, Africa, and Asiatic Russia, and exclude immigrants from the rest of Asia. We have found it necessary to prevent unfair competition and trade practices, to dissolve trusts, and to regulate the railroads. Market price as a guide to the production of educational services was found wanting, and so we have made education free to all, including in many places books and sometimes meals for school children. We have concluded as a nation that it is desirable to suppress the liquor traffic completely. City zoning laws substitute conscious city planning for the blind play of automatic forces. Employers are compelled to compensate workmen for industrial accidents, and some of our states have minimum-wage laws for minors.

But while we have departed widely from the principles of laissez-faire and are moving steadily in the direction of more and more social control of industry, it is nevertheless true that the automatic forces are still the dominant ones. They are the all-pervasive factors that regulate the billions of contractual relations of purchases and sales in our economic order.

Governmental control is important, but it does not enter every nook and crevice of our industrial life, as do the automatic regulative forces. Suppose we should attempt to abolish freedom of trade, private enterprise and competition, and substitute state control over each one of the billions of transactions that take place year in and year out in our complex industrial society. The task would be impossible. One has only to contemplate the magnitude of such an undertaking to realize the extent to which the automatic regulative principles control our modern economic life.

These principles may themselves be utilized in a scheme of social control. Consider the matter of accident prevention. We may seek to eliminate industrial accidents by a program of government regulation and inspection. Or we may make it financially worth while for employers following their own self-interest to eliminate accidents as far as possible. This the Workmen's Compensation Law does. The lower the accident rate the lower the premiums that have to be paid. This sort of social control which motivates the employer by appealing to his self-interest is far more effective than coercion, inspection, and the ordinary methods of law enforcement. We cannot expect that employers following their own self-interest in a régime of laissez-faire will conduct industry in such a way that desirable social conditions will ensue. But society can set up social arrangements which will induce private enterprise following its own self-interest to direct industry toward the ends conceived by society to be desirable. Thus, if society places the financial responsibility for industrial accidents and unemployment upon business men by compelling them to pay the costs of accidents and unemployment insurance, they will find it to their self-interest to buckle down to the task of eliminating these evils. Thus the principle of social control may itself utilize the automatic regulative forces. And even though nonpecuniary motives should sometime, in a less materialistic age than the present, become dominant, self-interest might still be the dominant motivating factor.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "Regulation implies *system* and *order.* A systematic or orderly world is one which is subject to the regulation of definite conditions and forces." Defend or oppose this statement.

2. "That economic activities are regulated is shown by the fact that a fairly definite proportion of our productive power is devoted each year to the production of wheat, shoes, cloth; that the balance of trade between the United States and England remains fairly constant from year to year; that the per cent of the value of manufactured goods paid out in wages is substantially the same year after year." Defend this statement.

The following table gives the percentage ratio of wages to value added by manufacture, in the United States:

Year	Per Cent
1899	41.5
1904	41.5
1909	40.2
1914	41.3
1919	42.1
1925	40.1

3. "A rate of wages determined by unrestricted individual competition is a 'natural' rate of wages; a rate of wages determined by combination or by legal enactment is an 'artificial' wage." Do you agree? Defend your position.

4. "Thus, strictly speaking, every development of social structure and function from the family to a police force, from the institution of personal property to the provision of public parks and libraries, from the primitive taboo to the most complicated act of Parliament, is alike 'artificial,' that is to say, the product of human intervention, the outcome of human activities. The plain truth is that to apply the antithesis of 'natural' and 'artificial' to social action is sheer nonsense." — BEATRICE POTTER WEBB. Do you agree? Why, or why not?

5. "We must remember, again, that if this system [the present economic order] really possesses the vaunted superiority of 'going by itself,' it is certainly not true that it made itself. If it goes by itself it is because the machinery is by this time well established. But when it had to be started — when individual ownership had to be created, with all its attributes of rent and interest and so forth — it required centuries of conquests, a hundred revolutions, a thousand laws, and all the might of kings, nobles, or parliaments. And, to tell the truth, this transformation still goes on without ceasing, so that it would be very hard to discover how much of the so-called natural order remains underneath the existing economic order." — GIDE. Argue for or against this proposition.

6. "Only free competition is right; co-operation [that is, co-operation between competitors] is, prima facie, wicked. Co-operation, in the sense indicated, is only a pleasant name for combining to take advantage of your customers." — TAYLOR. Argue for or against this proposition.

CHAPTER IV

THE ORGANIZATION OF PRODUCTION

The Function of the Entrepreneur. In this chapter we shall consider the function of the entrepreneur and the structural forms of organization through which entrepreneurship is carried on. By entrepreneur we mean the person or group of persons who control the business unit. If the business is a large one, several persons may jointly share in the function of entrepreneurship. At the outset it should be made clear that by control we do not mean management in the *narrower sense* of that term. Very likely the entrepreneurs employ a salaried manager to run the business. Such a manager is simply a paid employee and is in no sense an entrepreneur. He is working under authority and may be dismissed at any time or at least at such time as his contract terminates. The fact that many entrepreneurs manage their own business in no way implies that the details of the management are a part of the entrepreneurial function proper. In many small businesses the entrepreneur not only manages the business himself but also does a great deal of manual labor. Yet no one has ever contended that manual labor is a function of entrepreneurship. A hundred years ago, when all business was small-scale and almost all individual proprietors managed their own business, it was natural that management should be thought of as the essential part of the function of an entrepreneur. In modern society, with our large-scale production, division of labor, and high degree of specialization, the management function has split off in large part from the function of entrepreneurship. In many large-scale firms the entrepreneurs frequently do not manage the business at all; that function is delegated to a group of salaried officials selected and appointed by the entrepreneurs. Business managers frequently originate policies, but the entrepreneurs retain final control by their power to step in and alter the policy or oust the managers if necessary.

Yet in spite of the fact that it is necessary to distinguish between salaried managers and entrepreneurs, the point must not be overlooked that *control* implies *management*, in the *broader sense* of that term. The entrepreneur is the final coördinator of the factors employed, including even the salaried managers themselves. The going concern is under the final control of the entrepreneur, or group of entrepreneurs, and the ultimate aspects of management are therefore in their hands. In this sense it is quite correct to speak of management as being a function of entrepreneurship.

It is often said that the essential function of entrepreneurship is the assumption of risk. If this statement is intended to mean that entrepreneurs are the only persons who assume the risks of a business, we cannot agree. If the statement is merely intended to mean that the assumption of risk is inseparable from entrepreneurship, we must of course agree. But any human activity entails risks. A wage-earner assumes risks. He assumes the risk of accident and industrial disease. His employer may go into bankruptcy and may fail to pay any wages whatever. The hired manager assumes risks. He runs the risk of failure and loss of reputation, in which case he will not only lose his present position but will find it difficult to obtain another one. Those who lend money to the concern run the risk of possible loss of all or part of the funds advanced. Any person who does anything assumes risks. It is clear, therefore, that risk-taking is not a special characteristic of the function of the entrepreneur.

But does not the entrepreneur at any rate assume a certain special risk? His capital, it is said, is placed in the front-line trench. It withstands the first shock of loss; it serves as a buffer between possible losses and the claims of wage-earners, salaried workers, and bondholders. While this statement is, generally speaking, true, nevertheless our view is that it confuses the function of the stockholders with the function of the entrepreneur. To be sure, in corporate enterprises the entrepreneurs are generally stockholders, but not all stockholders are entrepreneurs. Nor is it necessarily true that entrepreneurs are stockholders even in corporate enterprises, though ordinarily they would be.

The entrepreneurs are the persons who are *in control* — who dictate the business policies of an enterprise. In theory this would include, in the case of a corporation, the entire body of stockholders, or at least all the holders of common stock, since all holders of common stock generally have voting power according to the number of shares held. But in actual practice the great body of stockholders in large corporations do not look upon their shares of stock as instruments of control at all. They seldom bother to vote their stock, they know nothing about the affairs of the company, and they exercise no voice whatever in determining its general policies. In short, they exercise no control whatever over the enterprise. They are in no sense entrepreneurs. The small group of stockholders who dominate the company, elect the directors and officers, and dictate the policies of the enterprise — these persons are the entrepreneurs.

Occasionally even this inner circle of stockholders loses its control. It may be that they have placed themselves in such a position financially that a banking house, to whom the company is under financial obligation, is able to step in and dictate the company's policies. It may be that the banking house can force the establishment of a voting trust which will insure a continuity of such business policies as will serve the interests of the bank until the obligations are met. In this case the bank would be the real entrepreneur.

We have used the phrase "dictate business policy." This of course does not mean that any group of entrepreneurs have unlimited arbitrary power over any business. The business policies in question are under their control within certain limitations. There are, first of all, the obvious limitations of the laws of the land. Then there are also such forces as custom, the demands of the public, the opposing wills of the organizations, groups, and individuals with whom they deal. Who, indeed, controls American industrial life, for example? It is altogether too simple an answer to say that it is controlled by some ten million entrepreneurs. Millions of these entrepreneurs — the farmers and small storekeepers, for example — have indeed no control over the broad currents of economic activity. In a sense they are inconsequential cogs in a great "commercial feudalism" controlled by powerful financial groups who are the real barons and kings in the modern world. And yet it is certainly naïve to think of modern industry being completely dominated by a few absolute and all-powerful financial monarchs. Shall we say that the wage-earning class, for example, have no control over industry in modern countries? Certainly they are a power to be reckoned with, and so are the farmers. Control is, in fact, in modern times a diffused thing. It shows itself in the control of property, in the ballot box, in trade unions, in trade associations, in farmers' organizations, in women's clubs, in public opinion. Control is much broader than entrepreneurship. Entrepreneurship refers to a special kind of control. It operates within a certain range, subject to the broad influence of social control.

Forms of Entrepreneurial Organization. We must now indicate in greater detail the various forms of business organization through which entrepreneurs function. We shall briefly describe five: (1) individual proprietorship, (2) partnerships, (3) corporations, (4) coöperative associations, and (5) political units, such as a municipality, state, or nation.

Individual proprietorship was the dominant form in the period of small-scale industry, and it is still of great importance wherever the small business unit survives. Agriculture is still nearly as important as manufacturing in America, and it is controlled almost exclusively by entrepreneurs who are individual proprietors. There are, all told, approximately six million of them. Retail stores are still operated to a considerable extent by individual entrepreneurs. This is especially true of business in small towns and of corner stores in the residence sections of large cities. In manufacturing there are still some hundred and forty thousand establishments operated by individual proprietors. In the case of stores and farms, especially, the individual proprietor is frequently entrepreneur, manager, capitalist, and manual worker, all in one.

The great bulk of economic activities which we classify under the head of manufacturing, transportation, and finance are carried on by partnerships or corporations, especially the latter. The *partnership* is an association of individuals engaging in a business. It is formed by means of a partnership agreement, which agreement is a contract between the individuals entering into the partnership. Generally, but not always, each member of the partnership furnishes a part of the capital and participates actively in the management of the business. At all events the entire property of each member is liable for all the debts of the firm. This is known as unlimited liability. Each partner assumes risks as great as though he alone were the proprietor of the business. To be sure, the risk is shared jointly by all the partners, and therefore the liabilities of the firm would devolve on any one partner only if all the other partners proved to be absolutely propertyless. In one respect the risk is greater than that of an individual proprietor, for any partner can enter into a contract relating to the business in question, which is binding upon the firm even though the other members have no knowledge of the transaction and have never given their consent.

Each and every partner may or may not be an entrepreneur. Entrepreneurship entails, let it be remembered, financial control. The silent partner who has an investment in the business but assumes no active control is not an entrepreneur. Only the partners who have investments in the business and actively control the final policies are, properly speaking, entrepreneurs.

The *business corporation* is also an association of persons. But it is a very special type of business association. The corporation is a creature of the state. A group of persons decide to form a corporation. They draw up articles of incorporation in which they state the name of the corporation, the purposes for which it is formed, the amount and kinds of stock that the company will issue, the location of the main office, the duration of the charter, and the names and addresses of the persons seeking to incorporate. These articles, properly drawn up, are submitted to some state officer, generally the secretary of state; and if they are legally drawn, he approves and files the application, which then becomes the charter of the newly created corporation. The state having granted the charter, the corporation is looked upon as a legal person, separate and distinct from the individual incorporators. Being a person in the eyes of the law, it can sue and be sued, own property, and enter into contracts. The shares of stock are transferable, and therefore the life of the corporation does not depend upon the life of the individual stockholders. All the stockholders might be killed simultaneously in an earthquake, but the stock would pass to their heirs and the corporation would live on. It thus has continuity of life separate and distinct from that of the constituent stockholders.

Unlike the partnership, which usually entails unlimited liability upon each individual partner, the individual stockholders who associate themselves together under the corporate form of organization usually escape all personal liability. As a general rule the creditors of the corporation cannot attach the individual property of the stockholders at all. But in the case of national banks, and usually in the case of state banks, the stockholders are liable for the debts of the corporation up to an amount equal to the par value of the stock held by them. In some states all stockholders are so liable except the stockholders of corporations engaged in manufacturing and mechanical pursuits. But at most the liability of stockholders is a limited one. The stockholders, of course, always run the risk of losing their investment in the corporation, should it encounter business difficulties and become insolvent.

The stockholders elect the board of directors, who in turn elect the officers and control the policies of the corporation. As a rule each stockholder has a vote for every share of stock owned. In theory, therefore, the corporation is democratically controlled by the entire body of stockholders. In practice, however, the great body of stockholders in large corporations do not exercise their voting power at all. Some of our large corporations have more than a hundred thousand stockholders. Each of these persons has a right to attend the annual stockholders' meeting at which the board of directors are elected and general policies decided upon. To the uninitiated it might be supposed that such stockholders' meetings are very large affairs, yet the usual rule is that only a few persons attend. These are the people who dominate the business or who represent those who dominate it. They are large stockholders who own a controlling percentage of the total stock or who have succeeded in getting a sufficient number of the general body of stockholders to sign over to them proxies (giving power to vote the stock in lieu of the owner) sufficient to control the situation. These active dominating persons who actually control the business are clearly the real entrepreneurs.

Stockholders who exercise no control over the business are not, properly speaking, entrepreneurs. Nor are minority stockholders who seek to control the business, but fail to do so, in any sense entrepreneurs. These stockholders are in reality in the position of investors or speculators who have put their money into a business over which they have no control. It is true that they expose their property to the same risks as do the entrepreneurs proper (that is, the controlling stockholders). But while entrepreneurs necessarily run financial risks, risk-taking alone, as we have learned, divorced from control does not constitute entrepreneurship.

Sometimes the bondholders become the entrepreneurs. Bondholders are persons who have lent money to the corporation, and the bonds are evidences of the loans made. Ordinarily they have no control over the affairs of the corporation. But they always have a contingent control. If the interest on the bonds is not paid, or if the principal is not paid when the bonds mature, the bondholders may foreclose on the property of the corporation and assume a position of active control of its policies. Here again certain dominant bondholders are likely to become the real entrepreneurs.

Another form of association which more and more assumes the responsibility of business ventures is the *coöperative society.* Two types are of chief importance: (1) consumers' coöperation and (2) marketing coöperation.

The first is an association of consumers who carry on a retail-store business. Several local consumers' coöperative societies frequently form a wholesale society which engages in a wholesale business, supplying the goods at wholesale to the various local stores. Such wholesale societies sometimes engage in certain manufacturing businesses, producing a portion of the goods supplied to the affiliated retail stores. Thus the British Wholesale Society manufactures soap, shoes, and clothing, owns a fleet of ships in which tea is shipped from their own plantations in Ceylon and wheat from their own wheat fields in Canada.

Marketing coöperative societies are associations of producers interested in marketing their products. The California Fruit Growers' Exchange, the American Cotton Growers' Exchange, potato exchanges, milk producers' associations, and farmers' elevators are examples. Coöperative creameries, while engaged chiefly in manufacturing, are also concerned with marketing.

Coöperative societies differ from corporations in several important respects. Profits are distributed on a patronage basis after a certain fixed dividend of 5 or 6 per cent has been paid on the stock. In the case of consumers' coöperative societies dividends are paid on the purchases made by each member, and in the case of marketing coöperation each producer is paid a dividend based on the volume of his products marketed through the associations.

Another important difference between the coöperative society and the corporation is in the matter of control. Coöperative societies are usually much more democratic in this matter than are corporations. In the first place, no shareholder in a coöperative society has more than one vote, regardless of the number of shares owned. In the second place, shareholders in coöperative societies exercise control over the business to a greater extent than do the rank and file of stockholders in a corporation. Entrepreneurship is therefore somewhat more widely diffused in the case of coöperative societies.

Finally, we must consider *political units in the rôle of entrepreneur.* Municipalities frequently own and operate their street railways, gas plants, electric-light plants, and waterworks. All western governments operate postal services, which are enormous business enterprises. On the continent of Europe certain of the governments own and operate the railroads. Some of our states have from time to time undertaken to operate flour mills, elevators, cement works, and have engaged in hail insurance, farm-mortgage banking, and commercial banking.

Who are the entrepreneurs in these cases? To say that they are the various political units involved is too general. Entrepreneurship is personal, and in any sort of business enterprise it should be possible to discover the persons who are the real entrepreneurs. In the case of democratic political units the voters are ultimately the ones who control the enterprise. From the immediate standpoint the political party or group of parties who are in control of the government dominate the situation. The party in turn is responsible to the majority of voters who elected it to office. Yet the minority party and its body of voters usually have large representation in the legislative branch of the government and are by no means without power. This very fact is one of the difficulties in the way of efficient governmental administration of business enterprises. The minority party is likely to do everything in its power to hamper and block the activities of the majority party in order to discredit it. In the corporation the majority absolutely and completely dominate the situation. If the minority do not like the way the dominant element run the enterprise, they can always sell their stock and get out. This situation gives rise to unified and harmonious control. But in the case of the government the defeated minority cannot get out, and they are always present to interfere with the carrying out of a clearly defined, consistent, and homogeneous policy. To be sure, even in the case of private corporations it sometimes happens that there are several groups of powerful stockholders, none of which can obtain a majority control. Compromise is then essential, and a clearly defined consistent policy becomes difficult and perhaps impossible.

But, it may be asked, are not the taxpayers the real entrepreneurs in government-owned enterprises? We have learned that entrepreneurship means financial control and that the assumption of risk is inevitably associated with such control. Although the taxpayers assume the risk, yet the voters have the final control. Here at least it would appear that entrepreneurship does not necessarily entail any assumption of risk.

From one point of view the distinction made is more imaginary than real. In all modern societies all self-supporting voters are taxpayers, though they may not always be aware of it. Every self-supporting individual, though he may not own any property, at least rents a room or a house. As tenant he indirectly pays the tax on the house. Tariff duties, consumption taxes, and so on are finally borne in the main by the general body of consumers, whether property owners or not. Yet in fact, in so far as modern governments are tending in the direction of abandoning indirect taxation and exempting low incomes from taxation, it would seem that control over government enterprises is being divorced from any assumption of risk. On the other hand, it may be argued that every citizen has an equity in the government-owned wealth and the total government income, and both would be affected by the success or failure of government-owned ventures.

The wider the dispersion of ownership and control of a business the more difficult it becomes to locate the entrepreneur. In the case of the individual-proprietorship business the case is clear and simple. So also in the case of the closed corporation, which is owned by a small group of men or even by a single family. But when we come to the great modern corporations with thousands of stockholders, — publicly owned corporations, so called, — the case is very different. And when the state becomes the unit of business organization, the entrepreneur becomes a very elusive person indeed.

Perhaps it may be asked, Is not the concept of entrepreneurship exclusively a private-property concept? Under socialism, with private enterprise eliminated, would not the function of entrepreneurship disappear? It must certainly be admitted that to undertake and control a business in a régime of private enterprise and free competition is quite a different thing from the functions performed by a monopolistic state engaged in an economic activity. It is quite possible that the two functions are so widely different that it is misleading to speak of a completely socialized state as carrying out the functions of an entrepreneur. Obviously the objection to using the term is less if the state controls only a few economic activities in an individualistic and competitive society.

SUPPLEMENTARY STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The entrepreneur is the agent who assumes responsibility in productive undertakings." — TAYLOR. Argue for or against this statement.

2. "The entrepreneur is therefore the point on which the whole of the economic mechanism turns. On him converge all the factors of production; and from him also, as we shall see, proceeds all the income. . . . The entrepreneur, therefore, is both the great employer and the great distributor." — GIDE. Comment.

3. "They [entrepreneurs] represent that class of people who make rather than take a job. From one standpoint, these are the only two important types of workers anywhere in the world. We have men who attend to a task after it is pointed out to them and others who create their own task and continue developing it as long as they live. Wage-earners generally speaking, though with notable exceptions, belong to the first group, and entrepreneurs to the second." — BOUCKE. Is this distinction sound? Does it properly define "entrepreneur"?

4. "He [the entrepreneur] has been well called a captain of industry, for he commands the industrial forces, and upon him more than any one else rests the responsibility of success or failure." — ELY. Do you agree?

5. "In the last analysis, then, the majority stockholders, who control the policy of the company, are the chief enterprisers in the case; although, sharing the risk and responsibility with other stockholders and officials, they share the profits also with them." — LE ROSSIGNOL. Argue for or against this statement.

6. Do you think it would be better to say merely that entrepreneurship is one of many forms of human activity of economic significance, than to set it up as a fourth or special factor of production?

CHAPTER V

COMBINATION OF THE AGENTS OF PRODUCTION

The Background of Social Institutions. We have noted in a previous chapter that all physical goods and all services consumed by man come ultimately from two sources: from man himself and from the environment external to man. There are, then, only two primary means of production, man and nature. But the extent to which man can utilize the resources of the region in which he lives is affected by his power to command these natural forces. It is a commonplace to say that the high level of well-being of the people of the United States in the present time is owing to the richness of the land that they inhabit. Yet the American Indian had exactly the same natural resources for his environment as we have. But he made little use of them in comparison with the extent to which we have utilized them. In part this difference is doubtless to be explained by the higher innate capacity of the white race for the type of civilization which prevails in the western world. But to an even greater extent it is owing to the fact that we possess a body of knowledge, derived from scientific discovery and invention, which enables us to command and control the resources latent in our environment. Even in Europe, where essentially the same racial groups now possess the land as in earlier centuries, production is much greater than it was two hundred years ago. Life is easier, and people have, on the average, many more goods to consume.

There is thus a third element or condition, in addition to nature and man, governing the productivity of a people. This condition is customarily referred to as "the state of the arts." But we must be careful not to think of the state of the arts as a third primary means of production. It is a powerful force in determining the average productive capacity of man and nature, but it is not a concrete, tangible, objective thing like man and nature. The state of the arts is the medium in which men operate and not the materials from which they produce. Stated less abstractly, the state of the arts determines the extent to which men are able to draw forth and turn to their own purposes the resources of the earth and of their own bodies.

Still another condition must be mentioned as an element in determining the productive capacity of man and nature. The nature of the social organization of a people, including the forms of government, legal systems, attitudes toward government, and much besides, also determines whether they will get much or little out of their environment. Earlier civilizations differed from our own quite as much in the extent to which men were able to coöperate as they did in the extent of scientific knowledge. They differed from our own in the ways in which individual initiative was stimulated or retarded. They differed also in the main objectives of human activity. But these conditions are no more means of production than is the state of the arts. Like the latter they affect productivity powerfully, but they are not concrete physical agents that can be used for making goods. They are qualifying circumstances that vary as time passes, and they act through the modification of man's behavior.

The Factors of Production. Production, as we have noted, is carried on by man and nature. Human beings perform three types of service, each of a different order. These three types of service are (1) labor, (2) entrepreneurship, and (3) saving. By labor we mean any human activity, either of body or of mind, that leads to or assists in the production of economic goods or the rendering of economic services. Entrepreneurship involves the assumption of the responsibility for carrying on and controlling the individual business enterprise and indirectly all productive activity in society. Saving makes possible the indirect, or roundabout, method of production. By this method capital equipment is first made, and this in turn is used to produce consumers' goods. Saving enables labor and entrepreneurship working with natural resources to produce capital goods.

The term "capital" is employed in the study of production in a special sense. It includes all man-made agents of production. It does not include natural resources or mere evidences of property rights, such as stocks and bonds. Capital consists of concrete material goods and not of legal claims. Consumers' capital consists of finished goods ready for consumption, such as houses, musical instruments, libraries, and furniture in the hands of consumers. Producers' capital consists of wealth that is still in process, such as raw material, or of aids in production, such as machinery, factory buildings, and the equipment of railroads. But the major part of all capital is used by producers. Hence we shall generally use the term to mean producers' capital, and when consumers' capital is meant that meaning will be clearly indicated.

The concrete items that compose producers' capital are not exactly alike. Some, like factory buildings, are very *durable*; and others, like coal, are *perishable*; some, like coal and metallic money, are *unspecialized* and can be used to aid in the production of a variety of goods, whereas others, such as spinning machines and blast furnaces, can be used at only one step in the production of one commodity or a closely related group of commodities.

We find then that there are four essentially different types of service-rendering objects which may be distinguished from the point of view of economics: labor, natural resources, capital, and the entrepreneur. In the present state of the arts all these factors are necessary for production, although it is possible to imagine a society so primitive that capital and entrepreneurs would be entirely absent.

Since all are necessary to carry on any productive process, it follows that productivity will be affected very greatly not only by the sum total of them but also by the relative amounts of each that are available for combination with all the others. In the evolution of industrial society many different combinations have been made. In sparsely settled regions much land has ordinarily been combined or used with little labor, labor being the scarce factor. Before the Industrial Revolution changed so profoundly the state of the arts, capital was of little importance in the combination of the factors. And with the development of large plants and giant corporations the number of independent entrepreneurs has decreased, relative to the total population. In farming alone, the independent entrepreneur continues to hold his own.

The Relative Scarcity of the Factors. The way in which the four factors of production are combined depends in a large measure upon the supply of each. For example, whatever capital equipment is in existence will certainly be utilized somehow in the long run unless it has become entirely obsolete, and the manner in which it will be combined with the other factors will depend a great deal upon its relative abundance or scarcity compared with the other factors.

Of the four factors of production, natural resources are limited in quantity by nature. Population is constantly growing, and from it springs an increasing supply of laborers, both manual and mental, and entrepreneurs. Capital is rapidly accumulating. It is true that during the nineteenth century the land supply was also increasing. The railroads opened up vast areas of land in North and South America, Australasia, Russia, Africa, and parts of Asia. This phenomenal exploitation of virgin natural resources is one of the most important characteristics of the last hundred years.

But certain individual countries long ago experienced a limitation in the supply of the natural factor in production. Thus England at the close of the eighteenth century and during the first quarter of the nineteenth century was experiencing grave difficulties because of the fact that while her population was increasing, her food-producing areas were already limited.

Because of these difficulties there developed a considerable public discussion which resulted in the gradual formulation of the principle known in economics as the law of diminishing returns.

The Law of Diminishing Returns. The following is a statement of this law as given by John Stuart Mill, the great English economist, writing in 1848. "After a certain, and not very advanced, stage in the progress of agriculture, it is the law of production from land, that in any given state of agricultural skill and knowledge, by increasing the labour, the produce is not increased in an equal degree; doubling the labour does not double the produce; or to express the same thing in other words, every increase of produce is obtained by a more than proportional increase in the application of labour to the land."

This is a statement of the law of diminishing returns applied to land. In fact, from a social point of view, the application of the law to land or natural resources is of primary importance.

Natural resources (land) are relatively limited in quantity. We may therefore call them a constant, or fixed, factor. The supply of laborers and entrepreneurs and capital is constantly increasing. We may therefore call them variable factors. Now every increase in the quantity of these variable factors applied to the relatively constant factor land will probably increase the product, but the increase in the total product will not (beyond a certain point) be in proportion to the increase in the quantity of variable factors applied to land. In other words, the average product per unit of variable factors declines after a certain point is reached. This is known as the point of diminishing average returns.

Let us illustrate this principle concretely. Suppose we take 300 acres of land which we will consider the fixed, or constant, factor. Suppose there is applied by way of experimentation an increasing quantity of variable factors to this plot of land. The results of this hypothetical series of experiments are shown in columns 3 and 4 of Table I.

TABLE I

1 FIXED FACTOR: 300 ACRES OF LAND	2 VARIABLE FACTORS: LABOR AND CAPITAL	3 PRODUCT TOTAL	4 AVERAGE PRODUCT PER UNIT OF VARIABLE FACTORS	5 MARGINAL PRODUCT: ADDITION TO PRODUCT DUE TO LAST ADDED INCREMENT OF VARIABLE FACTORS
Acres		*Bushels*	*Bushels*	*Bushels*
300	1 man with normal equipment	3,000	3,000	
300	2 men with normal equipment	7,000	3,500	4,000
300	3 men with normal equipment	10,800	3,600	3,800
300	4 men with normal equipment	14,000	3,500	3,200
300	5 men with normal equipment	16,000	3,200	2,000

Point of diminishing returns

It is evident from Table I that in this series of imaginary experiments the point of diminishing returns is reached when the labor of three men with necessary equipment is applied to 300 acres of land. From this it would follow that even if land were so abundant that it was a free good, men would not cultivate more than 100 acres per laborer; for if the work were spread out thinly over more land, the returns would yield a smaller product per laborer. If 300 acres per laborer were cultivated, the product would be only 3000 bushels; at 150 acres per laborer the product would be 3500 bushels; at 100 acres, 3600 bushels. So far intensive cultivation is more profitable than the extensive cultivation of a plot too large to permit proper attention and care. If, however, only 75 acres per laborer were used, although the per capita product would decline to 3500 bushels, the product per acre would continue to rise. If only 60 acres per laborer were utilized, the per capita product would be only 3200 bushels.

If, therefore, there were an abundance of free land, all of equal grade, not more than 100 acres per laborer would be used, no matter how small the population of the country. But suppose the population is so large that there are not 100 acres per laborer. It may be that there are only 75 acres per laborer. Under this situation it is obvious that the natural resources will have to be utilized beyond the point of diminishing returns.

In frontier communities it has no doubt frequently happened that labor and capital were spread out over so much land that it was not utilized even to the point of diminishing returns; or to put it another way, they were not utilized up to the point of highest per capita returns. This sort of cultivation was certainly uneconomical, but persons familiar with frontier agriculture can testify that such wasteful and ill-advised methods were used. As population increased, the point of diminishing returns was reached. As population became denser, cultivation became more intensive, more labor and capital were used per acre, and agriculture was pushed far beyond the point of diminishing returns.

None of the other factors is limited to the same extent as are natural resources. The numbers of entrepreneurs and of scientific experts are limited partly by biological heredity, partly by educational and industrial opportunities. Only a certain proportion of each generation is fit timber for entrepreneurship, for the higher grades of managerial labor, and for the highest type of scientific work. There is always more or less of a movement from the labor group to the entrepreneurial group or vice versa. It must be remembered that labor as we here use the term includes the highest-grade salaried worker as well as the more ordinary clerical worker and manual laborer. The margin of entrepreneurship, that is, the point of indifference as to whether one is a salaried worker or entrepreneur, is a continually shifting one. Superior entrepreneurs tend to bid natural resources, labor, and capital away from marginal entrepreneurs (that is, entrepreneurs who are now making no profits above wages and interest on invested capital), thus making it impossible for them to remain entrepreneurs. The less efficient entrepreneurs are squeezed out and pushed into the ranks of wage-earners or salaried workers. Thus a new margin of entrepreneurship is established.

The supply of entrepreneurs relative to the supply of laborers is therefore flexible. There is a constant ebb and flow between the wage and salaried group and the entrepreneurial group. If superior entrepreneurs are unable to utilize all the available factors of production, inferior entrepreneurs will come into the field of entrepreneurship from the ranks of the wage-earners or salaried workers. Thus the void can readily be filled. Accordingly the supply of entrepreneurs is as expansible as is the supply of labor.

With respect to the other factor of production — capital — it is astonishing how rapidly it accumulates in modern society. The last few centuries, with newly discovered undeveloped resources, improvements, and inventions, have opened up a field for capital investment so vast as to preclude seemingly any possibility of sufficiency of capital accumulation, yet the capital has been forthcoming year after year. In fact, the increased production consequent upon these discoveries and inventions is itself the source, in large part, of this capital accumulation.

From the broad social standpoint, then, land (natural resources) is more limited than the other factors of production. To be sure, none of the factors is unlimited and all are scarce; but capital, labor, and entrepreneurs are certainly at the present time increasing faster than natural resources. These factors must therefore be applied to land more and more intensely as it grows increasingly scarce. As we proceed with our study we shall see the important bearing of this principle of diminishing returns from natural resources upon economic rent.

wages, land values, value of farm products, population, immigration, taxation, and a host of other problems.

Agents of Production. The productive process has thus far been discussed from a rather broad social point of view. We have referred to the factors of production in general terms, as labor, natural resources, capital, and the entrepreneur. And the combination of these factors has been treated from the social point of view.

In the actual process of production as we find it in individual businesses and as it appears to the entrepreneur who directs a business, the factors of production are not so simply classified. The entrepreneur thinks in terms of common labor, machine operators, foremen, clerks, departmental managers, and the like, and not in terms of labor in general. When he considers a combination of productive forces, he often considers the economic advisability of adding or subtracting a kind or grade of labor to other kinds, as well as adding labor in general to capital and natural resources. Likewise he breaks up the general factor — capital — into many subclasses and combines power units with machine units, or both with raw materials. In the same way natural resources are classified and combined, not as a factor but as several different varieties.

When the factors of production are thus split up into various subclasses, we shall refer to them as *agents of production.* On page 61 is a schematic representation of the means of production and their relationship to the four factors and a few of the agents into which they are classified in the individual business.

The Best Combination of the Agents of Production. Under this head we shall treat the problem of production as it confronts the entrepreneur in the individual business establishment. This problem consists of two parts: First, in what proportions shall he employ the various types of labor, natural resources, and capital goods which are available in the market? Here the problem is to find the most profitable combination of the employed agents. Second, how large a quantity of these agents shall he employ in conjunction with

PRIMARY MEANS	FACTORS OF PRODUCTION	AGENTS OF PRODUCTION
Man	Entrepreneur	Varying sorts of entrepreneur
	Labor	Common workers Machine operators Semiskilled workers Skilled workers Clerical workers Foremen Executives
	Capital goods	Buildings Other structures Prime movers (engines, motors) Machines Means of transport Raw materials Finished goods awaiting consumption
Nature	Natural resources	Building land Roadways, waterways Mines Forests Fisheries Crop land Pasture land Recreation sites

himself as entrepreneur? This has to do with the problems of large-scale production and the size of the business unit. Taken together the two problems are concerned with the most profitable combination of all the agents of production, including entrepreneurship itself.

An entrepreneur finds a certain market value on each of the various agents which he may wish to employ. In a later chapter we shall show how these valuations are established, but for the present we shall accept them as established, without inquiring into the reason why. An entrepreneur finds that he must pay established standard wage rates for the various grades of labor that he wishes to employ; he must pay certain prices for capital equipment and raw materials; and if he wishes to rent or to buy natural resources, he must pay the going rentals or prices.

The problem of the best combination of the agents of production must be distinguished from the social law of diminishing returns. The latter deals with great groups of agents, and it assumes the physical supplies of the factors of production without reference to their cost. It simply says, If you have a limited land supply and if your population and capital increase, you will find that the average physical product per unit of these variable factors will increase at a less than proportionate rate unless the additions to population and capital happen to be accompanied by a change in the arts which makes available more natural resources or increases the efficiency of human labor.

The proportionality of agents refers to the internal management of a plant. It assumes that the prices of the agents have been set by the market outside the plant. It asks, How should these agents be combined to get the maximum money returns from them, present prices assumed? It is a narrower and somewhat different application of the wider and more general law of diminishing returns.

If we assume as fixed the various prices which must be paid for the agents of production, how many of each will the entrepreneur employ? Let us take a concrete illustration. Suppose that an entrepreneur has built a shoe factory of a certain size. The factory building and the land on which it stands we shall call the fixed or constant factor. To be sure, an adjoining piece of land might be purchased and an addition built to the factory building; but except for such additions, which would only be made at long intervals of time, if at all, the factory building and the land on which it stands form a constant, or fixed, factor. To this fixed factor variable factors may be added. Shoe machinery may be installed, labor may be employed, and raw materials purchased. At first the entrepreneur may install a relatively small number of machines and employ relatively little labor. Gradually, as he builds up a wider and wider clientele of customers, he will install more machinery and employ more labor. He will apply more and more variable agents to his fixed plant. What will be the effect of these successive combinations of agents upon his cost per unit of output?

Average Product and Marginal Product. Tables II and III illustrate the principles involved in this problem. Table II gives the different numbers of agents employed in each successive experiment with the resulting *total product, average product per unit of variable agents,* and *marginal product,* that is, the product added by the last unit of the variable agents. The problem is simplified by lumping all the variable agents together as though they were homogeneous. In reality, of course, the variable agents consist of a complex assortment of different kinds of machines, tools, raw materials, and scores of different kinds of specialized labor.

TABLE II

FIXED AGENTS	VARIABLE AGENTS	TOTAL PRODUCT	AVERAGE PRODUCT PER 10 VARIABLE AGENTS	MARGINAL PRODUCT PER 10 VARIABLE AGENTS
Factory building and entrepreneur	10	1000	1000	
	20	2200	1100	1200
	30	3375	1125	1175
	40	4440	1110	1065
	50	5420	1084	980
	60	6250	1042	830
	70	6800	971	550

From Table II it appears that at first as variable agents are added the total product increases more rapidly than the increase in the variable agent. In other words, the point of diminishing returns is not yet reached. This point is reached, however, when 30 units of variable agents are employed. From this point on, the increase in the total product is less than proportional to the increase in the variable agents. These points are brought out clearly in the average-product column. The last column gives the marginal product; that is to say, the product added by the last increment of variable agents. The second 10 units of variable agents added 1200 units to the product. The last 10 units of variable agents employed added 550 units of product to the total.

Cost per Unit of Output. The individual entrepreneur must pay the going rates of wages, interest, and rent. He must likewise pay the established market prices for materials and equipment. But the cost of the *finished product* is not definitely fixed. A variation in the combination of the agents will result in a variation in the amounts of finished goods produced and so alter the cost per unit of output.

How far will it pay the entrepreneur to continue to add variable agents to his fixed plant? It is impossible to answer this question without knowing what he will have to pay for each of the agents he employs. Suppose he pays a fixed rent of $500 per month for the plant; suppose he pays himself $500 per month as entrepreneur and general manager; and suppose he pays $100 per month for each unit of the variable factors. We can then determine the cost of each unit of output. These results are given in Table III.

TABLE III. COST PER UNIT OF OUTPUT

VARIABLE AGENTS	OUTPUT	FIXED PLANT AND ENTREPRENEUR	VARIABLE AGENTS	TOTAL COST PER UNIT OF OUTPUT
10	1000	$1.00	$1.00	$2.00
20	2200	0.45	0.91	1.36
30	3375	0.30	0.89	1.19
40	4440	0.23	0.90	1.13
50	5420	0.18	0.92	1.10
60	6250	0.16	0.96	1.12
70	6800	0.15	1.03	1.18

In a successful business the money derived from the gross sales must at least cover interest and depreciation charges on the fixed plant, as well as wages, cost of raw materials, interest on investment and machinery, and many other expenses arising from the use of the variable agents. The output must sell for at least enough to cover the expense of the fixed plant, the services of the entrepreneur, and the variable agents. With every increase in output the cost of the *fixed agents* becomes less and less per unit of output. On the other hand, the cost of the *variable agents* per unit of output will necessarily decrease up to the point of diminishing average returns and will increase beyond that point. So long as the product increases more, proportionally, than the increase in the variable agents, the variable cost per unit of output declines; and when the product increases less, proportionally, than the increase in the variable agents, the variable cost per unit of output will obviously rise.

Now the entrepreneur will not cease adding variable agents at the point of diminishing average returns. In our illustration he will, in fact, if he is thoroughly enlightened, continue adding variable agents until he employs 50 units of them. The reason for this becomes clear when we observe that as the entrepreneur increases his use of variable agents from 30 units to 50 units the increased cost of the *variable agents* per unit of output is more than offset by the decreased cost of the *fixed agents* per unit of output. This will be seen from an examination of the total-cost column in Table III. The total cost per unit of output is the lowest when 50 units of variable agents are employed and the output is 5420. Beyond this point the increase in the cost of the variable agents per unit of output is greater than the decrease in the cost of the fixed agents per unit of output, and the total cost rises. The entrepreneur will naturally wish to produce at the lowest possible unit cost. He will therefore, if he follows a perfectly enlightened policy, employ 50 units of variable agents.

Lowest Unit Cost for Marginal Entrepreneurs. According to Table III the entrepreneur can produce 5420 units of output at a cost of $1.10 per unit. This is the lowest total unit cost at arginal Firm

which he can operate, and in the long run the price at which the output sells in the market must be at least that high or he will be forced out of business. At that price he just covers all the costs he has incurred and receives for his own work the same return he could reasonably expect if he were employed as a manager by some other firm. When an entrepreneur is just able to cover his expenses in the manner described, he is called a marginal entrepreneur, or the firm which he manages is said to be a marginal firm. There are, however, in every line of business some entrepreneurs who are able to turn out goods at total unit costs more or less below the corresponding costs of the marginal entrepreneur. Sometimes they can secure a reduction of cost through lucky purchases of the agents of production below their going market values. But these fortunate bargains are not a permanent source of advantage. An entrepreneur who can permanently produce at a cost below that of the marginal firm must be able to do so because he can cause the variable agents to bring forth a larger physical product than they do in the plant of the marginal entrepreneur. In other words, the low-cost entrepreneur is able to make a better utilization of the agents than is made in the plant of the marginal entrepreneur.

Now we have shown that the marginal entrepreneur will employ more and more variable agents until he finally reaches the point of lowest cost per unit of output. Or to put it in another way, he will continue to employ more and more variable agents as long as they more than pay for themselves. If we revert to Table II,[1] we will notice that while the second 10 units of variable agents cost $1000, they added 1200 units to the product. At $1.10 per unit this will amount to $1320. Obviously these second 10 units of variable agents more than pay for themselves. The third 10 units of variable agents also cost $1000, and they added 1175 units or a value of $1293. The fourth 10 units of variable agents added a value of $1172; the fifth, $1078; and the sixth, $913. It is clear that our entrepreneur could not afford to employ the sixth 10 units, because they added a value less than their cost. If our entrepreneur had added the sixth 10 units of variable agents, he would therefore have raised his cost above the normal price. On this basis he could not afford to remain in the business. He would be driven out by those entrepreneurs who were wise enough to employ the variable agents only up to the point where they paid for themselves, or, in other words, the point of lowest total unit cost.

---

[1] The student can readily make this calculation by multiplying the units given in the last column by the price per unit, namely, $1.10.

Lowest Unit Cost for Superior Entrepreneurs. How about the superior entrepreneur? Will he use the same combination of agents as is used by the marginal entrepreneur? Will he employ the variable agents farther beyond the point of diminishing returns than the marginal entrepreneur does? Again let us take a concrete illustration. We must remember that the superior entrepreneur is able to get a larger output out of equivalent quantities of agents than is the marginal entrepreneur. Let us assume that he gets 10 per cent more. Since he gets a larger product from the agents employed than does the marginal entrepreneur, it follows that his total unit cost will be less for each application of the variable agents. He can therefore afford to employ the variable agents beyond the point of lowest total unit cost, since he is able to sell the product at the same price as the marginal entrepreneur. Like the marginal entrepreneur he will continue to employ the variable agents as long as they more than pay for themselves. Table IV shows the value of product added by each successive amount of variable agents. It will be remembered that the product is 10 per cent higher than for marginal entrepreneurs.

It will be seen from Table IV that the superior entrepreneur can afford to employ a total of 60 units of variable agents. Our marginal entrepreneur could afford to employ only 50 units of variable agents. The marginal entrepreneur could get only 830 units out of his sixth increment of variable agents: and since this yielded a value product of only $913 while the

TABLE IV

	VARIABLE AGENTS	TOTAL PRODUCT	PHYSICAL PRODUCT ADDED BY LAST INCREMENT OF VARIABLE AGENTS	VALUE OF PRODUCT ADDED BY LAST INCREMENT OF VARIABLE AGENTS	COST OF ADDED VARIABLE AGENTS	NET GAIN FROM EMPLOYMENT OF ADDED VARIABLE AGENTS	TOTAL PROFIT (+) OR LOSS (−)
1.82	10 . . .	1100					$ − 790
1.24	20 . . .	2420	1320	$1452[1]	$1000	$452	− 338
1.09	30 . . .	3713	1293	1422	1000	422	+ 84
1.02	40 . . .	4884	1171	1288	1000	288	+ 372
1.00	50 . . .	5962	1078	1186	1000	186	+ 558
1.02	60 . . .	6875	913	1004	1000	4	+ 563
	70 . . .	7480	605	666	1000	− 334	+ 228

cost of the added increment of agents was $1000, it would not pay him to employ the sixth increment of variable agents. But the superior entrepreneur could get 913 additional units out of the sixth increment of variable agents, and this would yield a value product greater than the cost. The superior entrepreneur may therefore be expected to employ the variable agents *beyond the point of lowest total unit cost,* since the largest volume of profits can be made only by extending production in the various plants until the marginal cost equals the selling price. At this point the total unit cost will be somewhat raised, but the higher cost will be more than offset by the larger volume of sales. So long as the marginal cost is below the selling price, every addition to output will add something to the total volume of profits. The net profits are largest, as the columns in Table IV show, at the point where the marginal cost is equal to the selling price.[2]

These principles can perhaps be illustrated more clearly by the aid of diagrams. In our tables we illustrated our principles by showing how the output is affected by large changes in the quantity of variable agents used. We jumped from 10 units of variable agents to 20 units, from 20 to 30, and so on. The tables therefore consist of discontinuous series. In the diagrams, however, the series are continuous, and we can therefore measure how much the cost is affected by minute changes in the output resulting from the application of an increasing quantity of variable agents.[1]

---

[1] The price per unit is $1.10.

[2] At this point the student should calculate the marginal unit cost. This he can do by dividing the cost of the added variable agents (5th column in Table IV) by the physical product added by the last increment of variable agents (3d column in Table IV). This calculation will show that the marginal unit cost is $1.10 ($1.095) at the point of highest total profit.

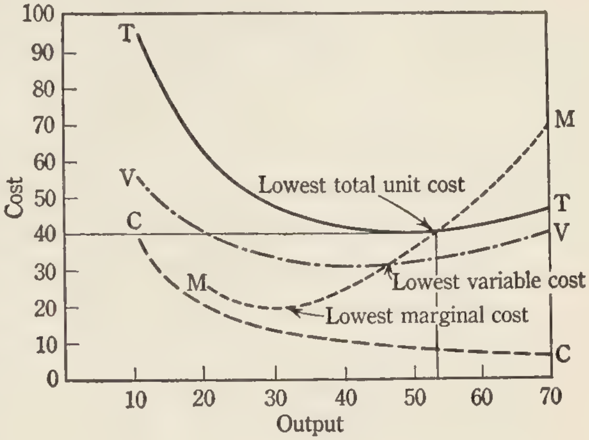

FIG. 2[2]

In Fig. 2 four cost curves are represented. *CC* represents the cost per unit of output in terms of the constant agents.

---

[1] See Alvin H. Hansen's "Prime Costs in the Business Cycle," *Journal of Political Economy*, February, 1924.

[2] The following table supplies the numerical data from which Fig. 2 is drawn:

A	B	C	D	E	F	G	H	I
Total Output	Total Constant Cost	Total Variable Cost	Total Cost	Constant Cost per Unit of Output	Variable Cost per Unit of Output	Total Cost per Unit of Output	Marginal Cost per 10 Units of Output	Marginal Cost per Unit of Output
10	400	550	950	40.0	55.0	95.0		
20	400	800	1200	20.0	40.0	60.0	250	25
30	400	990	1390	13.3	33.0	46.3	190	19
40	400	1240	1640	10.0	31.0	41.0	250	25
50	400	1600	2000	8.0	32.0	40.0	360	36
60	400	2100	2500	6.7	35.0	41.7	500	50
70	400	2800	3200	5.7	40.0	45.7	700	70

Columns A, B, and C are assumed. Column D is obtained by adding B and C. Column E is obtained by dividing B by A. Column F is derived by dividing C by A. Column G is obtained by adding E and F. The first item of column H is obtained by subtracting the first item of C from the second item of C; the second item of H is derived by subtracting the second item of C from the third item of C; and so on. Column I is the same as column H reduced from a 10-unit scale to a single-unit scale, on the assumption that if the marginal cost of 10 units is 250, the marginal cost of one unit will be 25; this is not strictly true, since the marginal cost of each of the 10 units is not the same, but the error involved is not great.

The larger the output the lower is the share of the constant cost attributable to any one unit of product. *VV* represents the cost per unit of output in terms of the variable agents. It will be noted that this curve at first falls until the point of diminishing average returns is reached, after which it rises. *TT* represents the total unit cost. It is derived by adding the per unit constant cost to the per unit variable cost for any given quantity of output. This can readily be seen by inspecting the diagram. At the output of 40 units, for example, the per unit constant cost is 10, while the per unit variable cost is 31. The total unit cost is therefore 41. Similarly at the output of 50 units the per unit constant cost is 8 and the per unit variable cost is 32, making the total unit cost 40. *MM* represents the marginal cost, or the added cost incurred in producing the last increment of product. An examination of the diagram will show that *MM* intersects *VV* at the point of lowest average cost of the variable agents; that is, at the point of diminishing average returns. This necessarily follows from the fact that so long as the marginal cost is below the average unit cost in terms of the variable agents, any addition to the output reduces the per unit variable cost. On the other hand, as soon as the marginal cost reaches the average variable cost, any further increase in output resulting from additional applications of the variable agents will raise the average variable costs. Therefore *MM* necessarily intersects *VV* at its lowest point. It will also be noticed that *MM* intersects *TT* at its lowest point. This necessarily follows from the fact that so long as the marginal cost is below the total unit cost, each addition to output lowers the total unit cost, whereas if the marginal cost exceeds the total unit cost, every addition to output increases the total unit cost.

As we have already shown, marginal firms may be expected to expand production under the pressure of competition to the point of lowest total unit cost, but no farther. This lowest total unit cost for marginal firms tends by competition to become the prevailing price. Superior firms will be able to produce at a lower total unit cost. But superior firms will tend to expand production beyond the point of lowest total unit cost. As was shown in Table IV, they can profitably expand production until the marginal cost equals the selling price. The result of this expanded production will be to raise the total unit cost somewhat, but the higher cost will be more than offset by the larger volume of business. In other words, it is clear that so long as the marginal cost is below the selling price, every addition to output will add something to the total volume of profits. Beyond this point it would be unprofitable to go, since the cost of the additional output would exceed its value. These points are illustrated in Fig. 3.

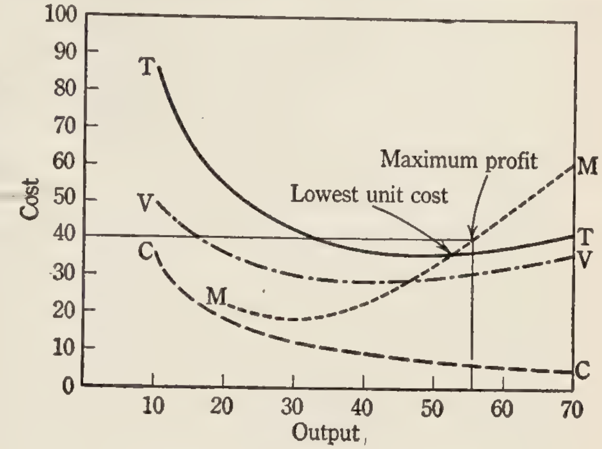

Fig. 3[1]

There will be a tendency, therefore, for both marginal and

---

[1] The following table supplies the numerical data from which Fig. 3 is drawn. For the manner in which these columns are obtained see footnote 2 on page 69.

A	B	C	D	E	F	G	H	I
Total Output	Total Constant Cost	Total Variable Cost	Total Cost	Constant Cost per Unit of Output	Variable Cost per Unit of Output	Total Cost per Unit of Output	Marginal Cost per 10 Units of Output	Marginal Cost per Unit of Output
10	360	500	860	36.0	50.0	86.0		
20	360	720	1080	18.0	36.0	54.0	220	22
30	360	900	1260	12.0	30.0	42.0	180	18
40	360	1120	1480	9.0	28.0	37.0	220	22
50	360	1450	1810	7.2	29.0	36.2	330	33
60	360	1920	2280	6.0	32.0	38.0	470	47
70	360	2520	2880	5.1	36.0	41.1	600	60

superior firms to continue to add variable agents until the marginal cost equals the selling price. For the marginal entrepreneurs this will coincide with the point of lowest total unit cost. For superior entrepreneurs this point will be beyond the lowest total unit cost. However, the total unit cost for superior entrepreneurs at the point of maximum profit will be more or less below the selling price, depending on the degree of superiority.

Our superior entrepreneur will no doubt be able to utilize effectively a larger fixed plant (either in the form of a larger factory or more factories) than the marginal entrepreneur can successfully control. In other words, he will not only utilize his fixed plant more intensively (by the employment of more variable agents) but he will also tend to utilize his own entrepreneurship more intensively. But entrepreneurial capacity is itself a limited factor; and as more and more plants are added, the point will at last be reached at which it will not be economical to take on any more plants. This leads us to the subject of large-scale production and large-scale management.

Large-Scale Production. We are now prepared to explain more fully the principles that determine the size of an establishment, to answer the question why large-scale production prevails in some industries while in other industries the small-scale plant prevails. We must distinguish sharply between two things that are likely to be confused: (1) the size of the firm; (2) the size of the plant. We may have two types of large firms: one, the firm owning many small establishments, such as the chain stores or the chain restaurants; the second, the firm owning one or more large-scale plants or establishments. There are two tests for determining whether a plant is operating on a large scale. It may be a large establishment because of the large amount of capital invested, though the number of wage-earners may be small. An example of this type is the modern giant flour mill, employing expensive equipment but relatively few wage-earners. On the other hand, if a large number of wage-earners are employed, the establishment will be considered a large one, even though relatively little capital is invested. An example may be found in the manufacture of cheap jewelry, such as decorative pins and collar buttons. Occasionally there are small plants, such as butter factories, that have no large total investment of capital but nevertheless a relatively large amount per laborer. Usually, however, plants that have a heavy investment per laborer are large-scale.

There are a number of questions that require answers: (1) What determines the size of the fixed equipment? (2) What determines the amount of labor and materials that will be utilized with this fixed equipment? (3) What determines the number of plants that will be brought under the management of a single firm? We shall consider each of these in the order named.

The size of the fixed equipment depends upon the extent of the market; the advance of the arts in that particular industry, — that is, inventions and technical processes; and the supply of available capital. The modern giant flour mill would be impossible in a society in which the market was limited to, say, a hundred thousand people. Even though the inventions had been made and large-scale processes were well known, it would not be economical to build such a huge plant unless the market were large enough to utilize fully the fixed equipment. It was improvements in transportation, particularly the railroad, that widened the market and made possible large-scale production. Granted the availability of an extensive market, the size of the fixed plant will depend in part upon the technical conditions prevailing in the industry. The market for farm products is world wide, yet farms remain relatively small because the technical conditions in the industry are such that small-scale farming is more economical. The same is true of cigar manufacturing — at least, of the better-grade cigars. In some industries the technical conditions are such that only a large-scale plant can succeed at all.

This is true, for example, in the steel industry. On the other hand, in the packing industry we have the huge packing plants at one extreme and the small country abattoir at the other. The large packing plant is based upon intensive division of labor and not upon technical processes requiring elaborate machinery. Finally, large-scale industry is impossible in a country which does not have sufficient capital.

The census of manufactures indicates than in many industries there is wide variation in the size of the establishments. This is partly due to the broad classifications used by the census. Thus all kinds of shoe factories are included under the heading "boots and shoes" — those producing high-grade custom-made shoes as well as those producing cheap standardized shoes. The former are made in smaller plants, the latter in large factories. But even where there are no such quality grades, there is sometimes wide variation in the size of the establishment. In some industries, however, small plants are nonexistent, being completely unable to compete. Table V gives the distribution of a number of industries as to size of establishments measured by the number of wage-earners employed. Only a few selected industries are given, in order to illustrate certain points. Group I includes industries in which large-scale establishments prevail. In these industries small establishments cannot survive. They are all industries that require much capital and a large investment in fixed plant. Group II includes industries that range all the way from small plants to huge establishments, but the large-scale establishments predominate. Small establishments survive, but they cut little figure in the total.

Group III includes industries which have all sorts of establishments, but the medium-sized establishment prevails. Group IV also includes industries that have establishments of all sizes, but the smaller establishments predominate. They are industries in which special machine processes of a large-scale sort are not greatly superior. Labor and management are more important than any special kind of capital equipment.

Group V includes those industries which have no establishments of huge size. Frequently they are industries which are tied down to local and restricted markets, or they are industries in which the technical processes are relatively simple, requiring little machinery, and in which labor or management or both are of much greater importance than capital.

TABLE V. NUMBER OF WAGE-EARNERS, IN THOUSANDS

	SMALL ESTABLISHMENTS (1–20 WAGE-EARNERS)	MEDIUM ESTABLISHMENTS (21–100 WAGE-EARNERS)	LARGE ESTABLISHMENTS (101–500 WAGE-EARNERS)	VERY LARGE ESTABLISHMENTS (OVER 500 WAGE-EARNERS)
Group I				
Automobiles	0.6	5.0	23.4	181.6
Cotton goods	1.1	17.9	146.0	265.9
Glass	0.3	7.8	42.0	27.4
Steel work and rolling mills	0.2	4.8	55.0	315.1
Watches	0.0	0.2	1.3	14.4
Group II				
Agricultural implements	1.6	5.2	14.7	32.9
Automobile bodies and parts	9.4	18.1	29.5	75.5
Boots and shoes	3.3	25.9	99.9	82.5
Railroad cars and repairs	3.0	27.4	150.7	303.3
Slaughtering and meat-packing	4.7	11.1	13.3	121.8
Soap	1.2	1.9	4.8	12.5
Tobacco	0.7	2.3	5.2	10.1
Group III				
Chewing gum	0.3	0.8	1.3	0.8
Men's clothing	24.7	61.1	51.0	36.4
Confectionery and ice cream	19.0	27.0	34.3	15.5
Foundry and machine-shop products	42.0	116.7	191.8	132.3
Furniture	10.2	50.8	57.8	19.5
Knit goods	6.8	35.3	80.4	50.1
Cigars and cigarettes	19.6	28.3	52.6	38.2
Group IV				
Bread and bakery products	66.7	26.8	31.6	16.4
Carriages and wagons	5.9	5.7	4.8	1.8
Flour-mill and grist-mill products	20.8	13.4	8.1	3.2
Women's clothing	45.8	86.0	31.4	2.4
Jewelry	7.8	13.2	8.8	1.1
Group V				
Automobile-repairing	43.8	8.3	3.0	0.0
Butter	10.0	5.6	2.0	0.0

We have discussed the conditions that determine the size of the fixed equipment. We turn now to our second problem. The amount of labor and materials that can profitably be employed in any marginal establishment will be that amount of labor and materials which will yield the lowest total cost per unit of output. This point depends fundamentally upon the law of diminishing returns, as explained in the first part of this chapter. Were it not for the fact that as more and more labor[1] is added to the fixed plant the point of diminishing product per unit of labor is reached, there would be no limit to the number of laborers that might be employed in any given plant. Because of the law of diminishing returns there is in every establishment a point of lowest total cost per unit of output, or it may be that there is a considerable range within which the total unit cost is approximately constant. This area or point of lowest total unit cost will generally be beyond the point of lowest unit cost in terms of labor alone. So long as the labor cost per unit of output rises less rapidly than the fixed capital costs per unit of output decline, — and they will of course decline with every increase in output, — the total unit cost will continue to fall. But the point of lowest total unit cost will vary with the price paid for labor, as compared with the capital costs. If wages are relatively low, more laborers can be added before the point of lowest total unit cost is reached. Let us revert for the moment to Table III. Suppose the cost of the variable agents is $25 per unit instead of $100; then the cost per unit of output in terms of the variable agents will be just one fourth as much as given in the table. The total unit cost will then be lowest when 60 units of variable agents are employed instead of 50, as given in the table. If, however, wages become relatively high compared with capital and other fixed or overhead costs, the point of lowest total unit cost will be reached sooner and less labor can profitably be employed in any given establishment.

---

[1] We are centering our attention here upon labor as the primary variable agent. There will of course be other variable agents also.

If for any reason the selling price of the commodity rises, more labor can profitably be employed. Thus during the war period, when prices were rising, establishments sought to increase output by hiring more labor. This was done even though there was no addition to the capital equipment. More labor could profitably be employed so long as the additional wage bill did not exceed the value of the product added by the more intensive utilization of the fixed equipment.

But the firm may continue to expand even though the plant has reached the most efficient size. It may build additional plants of equal size and efficiency. Certain advantages may follow from this procedure: (1) Freight charges may be considerably reduced by having plants located in different parts of the country. (2) Selling expenses may be reduced, for the greater the volume of sale the lower is the per unit cost of national advertising and of salesmanship. (3) A large firm can generally secure capital and credit more easily and on more favorable terms. (4) A large firm can frequently buy its materials on more advantageous terms. (5) A large firm can institute a single research establishment which will function for the several scattered plants. A striking example of this may be found in the case of the General Motors Company, whose research and experimental department is utilized for all the General Motors cars — the Chevrolet, Oldsmobile, Pontiac, Oakland, Buick, and Cadillac.

But there are limits to the economical expansion of a firm in any line of business. Management in a large firm is necessarily distributed through a hierarchy of officials. Orders are passed down through various grades of subordinates. The organization becomes unwieldy and top-heavy. Red tape supersedes personal supervision and attention. The large business unit faces at last the obstacle of the law of diminishing returns on entrepreneurship, and the point is at last reached at which it will be uneconomical to add another plant. Thus in most industries it is impossible for one business to expand until it occupies the whole field.

It must not be supposed that business men are always clearly aware as to what constitutes the most efficient size either of an individual plant or of the business as a whole. But that there is a most efficient size both of plant and of business unit cannot be doubted. It is a case of the survival of the fit. Many a business has been forced into bankruptcy by expanding beyond the limits of economy and efficiency. No doubt a business unit frequently continues to increase the size of its plant as long as the market is expanding. But such ill-considered growth will sooner or later lead to trouble. Such a firm is likely to find that smaller competitors with less top-heavy and unwieldy organizations are producing at lower cost. But it is not easy for businesses to stop growing when the most efficient size has been reached. The efficient concern is likely to grow; thus it may be big because it is efficient, and not vice versa. There is, however, danger of becoming too big. Hence businesses tend to pass from youth to maturity and thence to old age and final dissolution. Younger and more efficiently managed businesses take their places.

CONCLUSIONS

1. The law of diminishing returns affects (1) the entrepreneur's combination of employed agents, (2) the size of the business.

2. The most economical combination is beyond the point of diminishing average returns.

3. The marginal entrepreneur will add variable agents until the lowest total cost per unit of output is reached.

4. The superior entrepreneur will add variable agents beyond the point of lowest total cost per unit of output.

5. All entrepreneurs will continue to add variable agents as long as these added agents more than pay for themselves.

6. The size of a single plant depends upon (1) the market, (2) current technical processes, (3) the supply of capital.

7. The size of the firm depends upon the number of plants that can economically be operated under the directive control of a single management. Here also the law of diminishing returns operates.

PROBLEMS FOR DISCUSSION

1. In a period of intense business prosperity all firms tend to expand beyond the point of lowest total unit cost, but not beyond the point at which the marginal cost is equal to the selling price. Explain fully.

2. New businesses tend to be in a stage of decreasing cost; that is, an increase in output would result in lower total unit cost. Show why this is the case.

3. The railroad business as a rule tends to be in the stage of decreasing cost. What is the reason for this?

4. "We ought to force marginal entrepreneurs out of business in order that all our factors of production may be employed by superior entrepreneurs. Thus we would greatly increase the total national production." Criticize.

5. "The Ford Motor Company illustrates the point that there is no limit to the advantages to be gained from large-scale production." Criticize.

6. "The larger plants are bound to drive the smaller ones out of business sooner or later, and so in the end the tendency will be for one huge company to monopolize the field in each industry." Criticize.

7. "When the market is very small, no person can have any encouragement to dedicate himself entirely to one employment, for want of the power to exchange all that surplus part of the produce of his own labour, which is over and above his own consumption, for such part of the produce of other men's labour as he has occasion for." — ADAM SMITH. What bearing does this have upon large-scale production? What is the relation between division of labor and large-scale production?

CHAPTER VI

MARKETS AND THE MARKETING FUNCTION IN PRODUCTION

Definition and Classification of Markets. The term "market" is a familiar one whose general meaning requires no explanation. If, however, we examine closely the usage of business and of everyday conversation, we discover that the term includes three fairly distinct concepts: (1) A market may be thought of as a place where goods are bought and sold. (2) It is also thought of as the aggregation of interrelated organizations whose business it is to buy, sell, transport, and store a given sort of good. (3) Finally, a market is often spoken of as the sum of conditions or forces that play upon and determine prices. The first may be called the geographical, the second the structural, and the third the price-making (or "market-forces") concept of a market.

The geographical concept usually includes, in economic discussion, something more than the layman supposes. Wheat has a market in every hamlet where a wheat-buyer collects grain from farmers. But the geographical concept of the wheat market may include the entire world. Hence we frequently hear it said that Liverpool, which is the most important European port for buying and selling wheat, is a world market for wheat. By this statement it is meant that Liverpool, because of its importance, is the place on which the great dealers in wheat center their attention. Strictly speaking, however, there is no world market for wheat, but only a large number of related or linked markets which center in Liverpool. To prove this we have only to consider that wheat might sell at Liverpool on a certain day for the equivalent in

American money of two dollars a bushel, while in other parts of the world a variety of prices were paid, because of interference with freedom of trade by tariffs and because of costs of transportation, storage, and cleaning, which must necessarily be incurred before the grain could be delivered in England.

B The structural concept emphasizes the business organization of the market. A simple type of organization is exemplified by the markets for farm produce, which are still found in many cities and in which the growers of vegetables and small fruits daily offer for sale the products of their gardens. Usually very little physical equipment is required for this market — merely a shed or rough building in which the sellers may display their products. To this place the local grocers, and frequently even consumers, resort to lay in their daily supplies. There are no specialists, no accurate grading of products, and no service for keeping buyers and sellers informed as to the prices that are being paid for the same commodity in different parts of the building. Neither is there any service for reporting the amounts purchased at different prices each day and the amounts likely to be available on any succeeding day.

At the opposite extreme from the local farmers' market are the great produce exchanges for grain, cotton, and wool. In the United States wheat is purchased from the growers by a multitude of local buyers who clean, grade, store, and ship to larger buyers or warehousemen in the central markets, such as Kansas City, Chicago, and Minneapolis. The railroads carry the product and deliver it to the central markets, where it is purchased by millers, by speculators, or by large elevator companies. In these markets the sales are made in the grain exchanges by brokers who are specialists in dealing in wheat. Elaborate services are maintained for keeping both buyers and sellers informed as to the size of crops in all parts of the world, the amount of the visible — that is, readily available — supply, world demand, and current sales. These informational services are kept up by private agencies and by governments. The trading in any given exchange is so centralized that every trader can ascertain almost immediately how much of each grade of wheat is being bought as well as the price on each transaction.

The third concept of a market is much broader than either of the other two. It is also more abstract. The geographical area over which buyers and sellers are in close relationship is easily visualized. So, too, is the concept of the market as a group of interrelated business organizations dealing in the same commodity at the different stages of its progress from the primary producer to the consumer. But market forces are not always concrete things. They are conditions, facts, and relationships. Yet for the central problem of economics — price-making — this concept is more important than either of the other two. For example, when wheat sells for $1 a bushel in Chicago on a certain day, and six months later for $1.20, it would be futile to look to either the geography of the market or its structure for the explanation of the change in price. In practically every instance a variation of price is found to result from a variation of one or more of the forces which control price-making. Perhaps the rise will have been due to a report of unfavorable weather in Argentina, to a threat of war in Europe, or to better business conditions in the United States.

The Variability of Price-Making Forces. The forces which affect the making of price are not the same for all commodities. On the demand side we find that some commodities are wanted for human consumption in their existing forms; others are wanted as aids in production. For some the consumer has a relatively stable demand, and for others his demand is fickle. On the supply side there are likewise a large number of different forces. The supplies of agricultural commodities are affected by weather, by the prevalence or nonprevalence of pests and plant diseases, by the money return that the farmers can get from other crops, by the supply of farm labor, and so on. The supply of steel, on the other hand, is not affected by conditions governing biologic growth, but it is affected by the cost of coal, the efficiency of transport, and other conditions which either do not affect the supply of wheat or affect it to a smaller extent. Both wheat and steel differ from certain consumable commodities in that the demand for them is not affected very directly by fad, whim, and the vagaries of demand that appear in the market for consumers' goods.

As a rule commodities which can be easily graded and therefore easily described, such as wheat, can be sold over wide market areas, whereas those which cannot be so graded cannot be sold over so wide an area. Or at least the buyers and sellers cannot deal with each other by letter, telephone, and telegraph. Furthermore, the commodity which is sold over a wide market must, as has been said, be easily transportable, 2 and it must be something that enters into trade in 3 a multitude of ways or that is widely consumed. When a wide market exists, it pays those who deal in the commodity to specialize; and it is likewise profitable to build up systems of communication which put dealers over larger areas in touch with each other. Such markets are said to be well organized, and within them the goods dealt in tend to uniformity of price. For some goods, such as old dwellings, the market is, outside large cities, very poorly organized. Houses cannot be classified, and each buyer has a somewhat different notion of what constitutes desirable qualities.

Finally, we must distinguish markets in which both buyers and sellers are producers in the ordinary meaning of the word, or at least dealers, from markets in which only the sellers are producers or dealers and the buyers are consumers. The former are called *producers' markets* and the latter *consumers' markets*. In the latter only consumable commodities are sold, whereas in the former the consumable commodities as well as raw materials and technical products are dealt in.

The contrast between producers' and consumers' markets is great even when we limit our examination to consumable goods. Producers' markets are usually better organized, the buyers have more competence to judge the quality of the things they buy, the force of competition is stronger, and the force of custom and habit is weaker than where the consumer is the buyer. The consumer buys for his personal satisfaction and for the satisfaction of his immediate family; his only criterion for judging whether a good is worth the price asked for it, when this price is common throughout the market, is his subjective estimate of the satisfaction he and his family are likely to derive from it. Although consumption, and not production, is rightly said to be the ultimate object of all economic activity, the art of choosing consumption goods intelligently and of utilizing them efficiently is, by contrast with business activities, extremely undeveloped. The buyer in the consumers' market is not an expert, and he cannot have the same knowledge either of the quality and serviceability of goods or of competitive prices as the dealer from whom he buys. Hence the advantage in bargaining is mostly with the latter. In the producers' market, on the other hand, the buyer and the seller stand on a nearly equal footing in bargaining power and in knowledge of facts relevant to the transactions into which they enter.

Hence it is not surprising to find that competition works more efficiently in the producers' markets and that uniformity of price for goods of equal serviceability to the dealer is much more common than is uniformity of price for equal goods in the consumers' market. If one goes from shop to shop in the retail markets in a large city, he will usually find some variation in the prices at which even standardized articles are sold. And if he turns to articles whose demand is greatly affected by style, the variations become greater. To a certain extent each store has its own customers, its own special demand. Each is a somewhat separate market for goods plus the services that ordinarily go with them.

Yet the lack of price uniformity in the consumers' market must not be overstated. Each of these partially distinct markets is related to other markets. One store cannot go too far above the prices of its competitors, or it will lose its patronage. A "most common," or ruling, price for standardized goods prevails, and even for style goods the variations are not extremely large during the season when they are in style. There are always some buyers who have the knowledge of quality, the time and inclination to shop, and the desire to find bargains, which enable them to ferret out the "best buys." They shift from store to store and spread the news of their bargains among their friends. Thus the seller is compelled to keep his prices nearly in line with those of his competitors. This fact is so well recognized that many merchants systematically "shop" competing stores.

The phenomenon of related markets is not confined to competition among the merchants in the same city. Recent investigations by Professor Secrist[1] have shown that the prices of men's ready-to-wear clothing in different cities over a large territory are related. The consumer is not restricted to the stores in the city where he lives. He knows about prices and qualities in other places through personal observation, through reports of associates, and through advertising. If he believes that prices are too high or goods not in style in the local shops, he can and does go to other towns or cities to buy. And since any group of buyers may have the opportunity to buy in several cities, stores in different places tend to be connected, or "linked," markets in somewhat the same way that the wheat markets of the world are linked markets. But the connection is far looser in the case of retail markets, and in consequence the uniformity of price is far less.

It would be an error to suppose that the tendency of prices in different, but connected, consumers' markets to conform to a most common price is solely the result of the activities of the relatively small body of consumers who shop widely. Much of the uniformity that exists comes from the pressure of competition among the dealers themselves. A storekeeper is not compelled to remain always in one location or to keep to one line of merchandise. If prices rise in one place or in one line, and if profits rise with them, competition through the establishment of rivals is likely to put in an appearance. New shops will be set up by men just starting as independent entrepreneurs or by those who are leaving one line in the hope of bettering themselves in another. Many groups of grocerymen have found in recent years that the maintenance of high prices or poor service in a given neighborhood creates a favorable situation for the establishment of a branch of a chain-store organization.

---

[1] *The Widening Retail Market.* A. W. Shaw Company, Chicago.

Yet in spite of the forces that tend to equalize the position of buyer and seller and to bring about uniformity of price the operation of the price-making forces is different in the consumers' market from what it is in the producers' market. And to some extent the forces themselves are different. Convenience, pleasant surroundings and pleasant personal relations, custom, habit, and indifference to standards of efficiency are much more potent in consumers' than in producers' markets. Moreover, the consumer must buy for personal use and rely upon his own subjective estimates for testing the satisfactions he will derive from goods sold at different prices. The buyer in the producers' market has much more objective criteria.

Marketing Functions. In economic writings of a generation ago very little attention was paid to the subject of marketing. The reader was usually told that "exchange" was an important division of economics, that the objective value of goods was determined by exchange, that the people who engaged in bringing goods from farms and factories to consumers were productive laborers, and that those who took part in marketing were called by various names, such as "wholesaler," "broker," "retailer," and the like, to indicate that they performed different functions. It is only within the past twenty years that the work done by marketing agencies has received the attention its importance warrants.

The present economic order is an exchange economy. No individual supplies himself with all the goods and services he consumes. Even the farmer, whose land often produces much of the food and fuel his family consumes, is compelled to sell to others a major portion of his products and to buy from others most of the consumable goods his standard of living requires. Among city dwellers self-sufficiency is still more impossible than it is on the farm. The manufacturer and his employees can use very little of their products, and professional workers require very little of their own services. All producers — farmers, manufacturers, factory hands, and professional workers — must rely on the aid of the people engaged in marketing to assist them in disposing of their products and to enable them to obtain the goods they consume. Accordingly, before we proceed further with the study of economic principles, it is very important to understand the functions performed by those who engage in marketing. The following classification, given by Professor F. E. Clark,[1] presents in compact form a general view of the functions performed by those who engage in marketing.

THE MARKETING FUNCTIONS

A. Functions of Exchange
  1. Demand creation (selling)
  2. Assembly (buying)

B. Functions of Physical Supply
  3. Transportation
  4. Storage

C. Auxiliary or Facilitating Functions
  5. Financing
  6. Risk-taking
  7. Standardization

The function of demand creation involves the work of finding buyers for the products of the farmer and the manufacturer. Goods do not sell themselves any more than plants grow without cultivation, or factories run without human direction and control. Even such staples as elementary raw materials and foods must be brought to the attention of the buyers, who must be informed where the articles are to be had, what their qualities are, and at what prices they are held. Luxuries and semiluxuries — the vast range of things used nowadays without any reference to physical necessity — require more effort and expense for selling. Not only must they be called to the attention of the buyers, but their uses must be explained, and either by advertising or by more direct methods the buyer must be convinced that they will satisfy a particular want. Sometimes these sales methods work to the detriment of the consumer, who may be persuaded to buy something which he finds later does not satisfy his needs or which is inferior to some other good he might have purchased at the same price. On the other hand, the selling function unquestionably extends the range of consumption and increases the satisfactions of the consuming public.

---

[1] F. E. Clark, *Principles of Marketing*, p. 11. Copyright, 1922, by The Macmillan Company. Reprinted by permission.

The function of assembly does not mean the work of gathering together the physical items in a stock of goods, for that is transportation. "Assembly" means that the middleman acquires the legal title to the goods, or at least the right to buy or sell them. He may have at his disposal, for example, a million bushels of wheat. A part he may actually own and have stored in various places. Another part he may not own at all, but may merely have agreements with other dealers to deliver it to him at a future date. But in either case he can control that amount — can decide to sell it or to withhold it from the market. He puts himself in a position to sell large or small amounts, as the buyer may wish, and he prepares to sell different grades to different buyers. In the marketing of manufactured goods the function of assembly includes a slightly different sort of work. The department store must stand ready to sell to its patrons literally hundreds of different wares, very few of which can be obtained from a single manufacturer. Someone must find out where all these wares can be obtained, the quantities in which they can be obtained, and their prices. Then he must offer them for sale in definite amounts. All this the large department store does for its patrons. The smaller store, which cannot afford to perform this function, usually relies on a wholesaler or some other intermediary.

The function of physical supply, which includes storage and transportation, is almost self-explanatory. Everyone knows that most of the articles he consumes are not produced at his door and that they must be brought from more or less distant places before he can consume them. He also knows that many products, such as agricultural crops, are harvested only at certain seasons and that they must be held in storage so that consumption may be continuous. Manufacturing is not ordinarily so much controlled by the seasons as is agriculture, but the demand for manufactured products may be seasonal. Hence storage must be provided if production is to be continuous. Furthermore, both production and consumption may be continuous and storage still be necessary because a stock must be maintained to give the consumer a choice among the goods he buys. The merchant is also compelled to keep a stock on hand because he cannot foresee exactly what the consumer desires. Often the latter makes his decision by looking over the different styles or varieties of goods displayed. Finally, local storage is sometimes resorted to when it is cheaper to ship goods to a particular city in large amounts and then store them until they are gradually consumed than it is to ship in relatively smaller amounts and avoid the major part of the expense of storage.

The auxiliary functions are so named because they do not directly assist either in the physical supply of goods or in the finding of buyers for them. Financing is a necessary function because time must elapse between the completion of the physical process of production and the sale of the goods to the consumers. The manufacturer and the farmer or some middleman must have money tied up in the stocks waiting to be consumed. The interest on these funds constitutes the financing service for which someone must be paid. Risk in marketing arises from (1) the possibility of an adverse change in price while the goods are in the hands of the dealer, owing to a falling off of demand or to oversupply, (2) the possibility of physical deterioration, and (3) the risks arising from a failure of buyers to live up to their agreements, either their failure to take goods ordered or their failure to pay for them. For many sorts of goods, particularly things for personal use, the risk of an adverse change in price is great. The style may change before the merchant has sold his stock, or the fad for a particular kind of good may not appear in one city even when it has appeared in other places. Or the season may be unusual, and the seasonal goods which the producer and the marketing agencies had reasonably anticipated would be in demand fail to attract buyers — witness the case of men's straw hats when the summer is cool and rainy. All these risks must be borne by the marketing agencies; and because the existence of a risk implies that there will be losses, the goods actually sold must bear the cost of those remaining unsold or sold at very low prices.

Standardization of products, which means classification into established groups or grades, is an important marketing function for some kinds of products. When the buyers and sellers cannot come together, with the goods before them, grading becomes extremely important. At present wheat is sold by description — that is, by grade — all over the world. A miller in England can buy "No. 1 dark northern spring wheat" and know approximately what he will receive, because wheat has been graded and standardized. Many commodities cannot be standardized, because technical tests for certain qualities have not been devised or because the make-up of the good is constantly changing. For both reasons it appears to be impossible to standardize style goods — clothing, millinery, shoes, hats, and so on.

Marketing and Specialization. In the earlier stage of economic development which preceded the present highly specialized economy each producer supplied a relatively limited area and relatively few consumers. Lack of cheap transportation, the slowness of communication, the simplicity of production, the absence of large plants, and political conditions all contributed to keeping the products of farmers and craftsmen from moving very far from the places where they originated. The producer could readily ascertain the kind and quantity of product his patrons required, and the consumer could more easily than now hold the producer responsible for poor quality and misrepresentation. By reason of the influence of custom and the close adjustment of production to needs prices were much more stable. There were, of course, exceptions to the conditions just sketched. A few commodities were imported into every country at extremely high costs for transportation, financing, and risk-bearing. But local production for local consumption was the rule.

Today conditions are very different. The average laborer's family consumes in the course of a year things that have literally come from every part of the globe: coffee, sugar, rice, rubber, and so forth. Even the products grown and manufactured within his own country are derived from the factories and farms of people he has never seen and knows nothing about, and most of these products have passed through the hands of many dealers, after leaving the factory or the farm, before they are delivered to him. Marketing has come to play a vastly more important rôle. And with this development some of the advantages of the old order have been lost. The relative cost of marketing has doubtless become greater in comparison with the cost of physical production; that is, the growing, mining, and processing of goods. In Table VI, p. 92 (taken from the Report of the Joint Commission of Agricultural Inquiry), are given estimates of the cost of selling at retail five important types of consumers' goods.

The percentages given in the table on the following page show only the *retail* cost of marketing, and not the proportion of the consumer's dollar which goes to pay for all the marketing services performed in respect to the five types of commodity.

TABLE VI. COST OF OPERATING RETAIL STORES, BY PERCENTAGE OF EACH DOLLAR OF SALES[1]

	1913	1916	1918	1920	1921
Clothing	25.90	25.20	26.50	25.70	24.03
Hardware	20.13	18.60	18.52	19.50	24.90
Shoes	24.14	24.45	24.07	25.25	26.88
Dry goods	26.90	26.30	27.30	26.00	No data
Groceries	14.70	15.20	14.30	14.60	16.80

A quantity of groceries, for example, which cost the retailer a dollar was bought by the wholesaler for about ninety cents during the years for which the table shows estimates.[2] The manufacturer of these foods had also incurred some selling expense, for which he was paid, and the dealers from whom he bought his raw materials were also compensated for their services. The congressional commission from whose report Table VI was taken is authority for the statement that "it costs more to sell and deliver the marketable products into which agricultural commodities are converted than to produce and manufacture them."[3] This estimate seems, on general observation, somewhat higher than would be true for all goods, but we must remember that many of the goods and services for which we spend our money incomes are not subject to such heavy marketing expense. Services rendered by the government, such as sanitation, education, and recreation, have few or no marketing costs. Personal services furnished by physicians, household servants, caterers, barbers, and the like are not "marketed" in the proper sense of the term, and therefore their prices include little or no marketing expense.

Another way of showing the importance and the cost of marketing is to compare the number of persons engaged in trade and transportation with the total number of gainfully employed in the United States. In 1920 the Census of Occupations reported over 41,614,000 persons gainfully employed. Of these, 4,423,000 were employed in trade and 3,064,000 in transportation. If we exclude from the classification of trade such doubtful groups as bankers and bank employees, real-estate agents, undertakers, and insurance agents and officials, there still remain 3,773,000, or 9.1 per cent of the total of gainfully employed, who were chiefly engaged in marketing commodities.[1]

---

[1] Report of the Joint Commission of Agricultural Inquiry, Part IV (Marketing and Distribution), p. 176.

[2] Ibid. p. 158.

[3] Ibid. p. 115.

The census classification of occupations is in some respects not indicative of the true industrial allegiance of the different groups of the population, and the methods of collecting census data leave considerable opportunity for listing people in occupational groups to which they do not belong. Statisticians who have reworked these data and supplemented them with other sources of information have probably arrived at a more nearly correct estimate of the numbers of the population in each occupational group. Mr. W. I. King finds that in 1921 there were 40,819,000 persons gainfully employed in the United States. Of these, 4,728,000, or 11.6 per cent, were in mercantile occupations, and 3,524,000, or 8.6 per cent, in transportation.[2]

Still another method of showing the cost of the present marketing system in general terms is to estimate the proportion of the national income received by those who derive their incomes from marketing. According to King the total current income of the United States in 1921 was $62,736,000,000. Of this total, $8,919,000,000, or 14.2 per cent, was received by the people in mercantile occupations, and $6,135,000,000, or 9.8 per cent, by those engaged in transportation.[3] Since a part of the business of transportation consists in carrying people and not goods, by no means all the income received by this occupational group can be counted as a cost of marketing. On the other hand, not all the persons concerned in marketing are included within the two groups for which the statistics have been given. Some of the workers connected with every manufacturing establishment and all farmers assist to some extent in marketing, and the incomes which these groups receive on account of that work is a part of the cost of the nation's marketing service.

---

[1] Fourteenth Census, Vol. IV, pp. 40–41.

[2] *Income in the Various States*, pp. 21–23. National Bureau of Economic Research.

[3] Ibid. pp. 32–33.

Against these costs which attach to the present system of marketing and which have probably become proportionately larger as markets have widened and grown more complex must be set very important gains. Although no statistical data are available to prove the point, there can be no doubt that the factory system, specialization, and large-scale production, all of which depend on wide markets for their existence, have increased tremendously the average productivity and income of the population of the United States and western Europe during the last century and a half. And if we should discard the present arrangements, which are admittedly expensive, and revert to the simpler market organization prevailing in the early eighteenth century, we should lose a large portion of our present income. Without the present system many kinds of goods could not be produced at all, and others only at higher costs.

PROBLEMS FOR DISCUSSION

1. If the chain of middlemen that connects the producer with the consumer were unnecessary, it would never have come into existence. Is this statement correct? Does it contain complete proof that present arrangements are the best that can be devised?

2. It is sometimes said that advertising and other forms of demand creation are entirely wasteful. Like modern armaments they result from destructive competition. If dealers did not attempt to sell goods, how would the consumer find new goods? Would he be able to discover as easily as now the cheapest place to buy? If it is assumed that advertising is purely competitive, is it socially wasteful?

3. In which kind of market should you expect prices to be more nearly uniform, producers' or consumers'? Why?

4. "Instead of having highly centralized live-stock markets in Chicago, South St. Paul, Omaha, and other places, we ought to have a packing plant within a radius of fifty miles of every farmer so that each farmer could sell his live stock directly to the packers." Discuss.

5. "A farmers' and wage-earners' coöperative marketing association ought to be organized so that farmers could sell their products directly to the consumers in the cities. In this way farmers would get more for their products and wage-earners would pay lower prices for foodstuffs." Discuss.

6. "These utopian reformers want to accomplish the impossible. They want everyone to sell dear, and at the same time they want everyone to buy cheap. They want a high price for beef on the hoof and a low price for beef on the table." Discuss.

BOOK II. VALUE

CHAPTER VII

THE PRICE-MAKING PROCESS: DEMAND AND SUPPLY

Price and Value. The value of any commodity is its power to command other goods in exchange. This value is measured by the quantity of the other goods for which the commodity in question exchanges. If shoes are bartered for cloth at the rate of one pair of shoes for six yards of cloth, then the value of cloth is measured in terms of shoes, and the value of shoes in terms of cloth. This is a correct representation of power in exchange; but it is not of much significance, because it only states the exchange value of each good in terms of one other good. If we wished to know the general exchange value of shoes, we should have to discover the rates at which they exchanged for all the other goods for which they might be bartered. To find these exchange relations would be practically impossible. And when exchanges were carried on by barter, it is hardly conceivable that anyone had very definite notions of the value of any good in terms of more than a few other commodities.

In present-day industrial societies barter has been superseded by the exchange of goods for money. The amount of money for which a unit of any given good exchanges is called its *price*. Since exchanges are now commonly made in money, it is possible to find the value of any commodity by comparing its price with the prices of a large number of other goods. If it is found that shoes sell for six dollars a pair and cloth for one dollar a yard, then it is easy to show that the exchange value of cloth and shoes at current prices is six yards of cloth for one pair of shoes. Comparisons of the price of cloth or shoes with the prices of many other goods would establish a still more complete measurement of the exchange values of both. And this comparison is constantly being made in many places and by many people. Hence we infer that the price of a commodity is a more or less accurate measure of its exchange value.

Price is not, however, an entirely accurate measure of value, because, as will be shown later, the value of money, which serves as a basis for the computation of the values of other goods, may itself change. But for the purpose of simplifying the analysis of the forces that control value we can assume that the value of money is constant and that if any good rises or falls in price it also rises or falls in value in the same proportion.

The Analysis of Price. If one should attempt to find out why a specific commodity sold for a certain price at a particular time, he would at once be confronted by a bewildering mass of facts whose relations are extremely complex. Moreover, if he tried to measure the relation of each fact to every other, he would never arrive at any conclusion, because the human mind is not capable of dealing with a very large number of facts or relationships at one time. Since he could not deal with the large number of facts and all their interrelationships, he would in all probability try to classify many of the facts into a limited number of groups, and to each group he would then attach a name. Thereafter when any fact relevant to his investigation was found, it would be immediately assigned to its proper group. This procedure, which is called classification, is common to all scientific investigation.

When different facts are lumped together, it is evident that some things will be put into the same group and treated as though they were exactly alike when, in reality, they are more or less different. This involves a certain amount of error, not only in the classification itself but also in any statement concerning the relationships between groups of facts. Hence the common saying that all generalizations — that is, all statements concerning the relation of one group of facts to another — are partly false. The greater the differences between the facts that are put into the same group the greater is the probability that any general statement about their causes and effects will be partly untrue. The danger of making erroneous statements because of the classifying of unlike things together is a very real one. It would be an error in reasoning, in many cases, to conclude that because the average wages of factory workers in any community had risen, the wages of any particular group of workers had also risen. Or again, when the general level of prices has risen by 10 per cent owing to a decline in the value of money, it does not follow that the price of each kind of good has risen by the same percentage.

Still another useful procedure in dealing with very complex relationships between facts is to rule out of consideration in the beginning all but the most important factors which appear to produce a given result. This leads to piecemeal analysis. The student may say to himself: Let us see how any given price would be affected by a few of the conditions that appear to control it. If he can discover that much, he is then prepared to take into account more facts, more "forces," until his analysis may approach a degree of completeness that will enable him to explain satisfactorily the price in question. This procedure is called abstraction. Like classification the method of abstraction is replete with opportunities for making false steps in reasoning. It might be concluded, for example, that if a larger crop of cotton was harvested, the price would be lower. But this deduction assumes that "other things are equal"; that is, that the demand of consumers for cotton cloth has not changed, that there have been no changes in the competition among the mills that manufacture cotton, that the growers will not be able to hold back part of the crop for future sales, and that the amount of cotton and cotton cloth remaining from last year's output is normal.

In the analysis of price we shall proceed in the following manner: We shall begin with the examination of what is called in economics "market price" and attempt to understand the way in which economic forces control prices for short periods of time. Having explained how prices are determined for a short period of time, we shall then go on to examine the forces that determine prices over a considerable period of time. This leads to the analysis of normal price.

Market Price. It is commonly said that market price is the price for which a good actually sells in any given market at any particular time. Yet in actual practice economists or market analysts seldom attempt to explain *particular prices.* The forces which cause a good to sell now at a little higher and later at a little lower price, and which give rise to minor differences in the price of the same good in the same market, are numerous and intangible. Moreover, small mistakes in judgment by buyers and sellers are too frequent to permit of the explanation of all the minor changes and discrepancies actually encountered in the market. The market price which we shall try to explain is that which is most common, or typical. If the price at which the bulk of the transactions in cloth take place during a certain day or week is one dollar, then we shall assume that this is the market price for cloth during that period. Many sales will have been made at somewhat higher or lower prices, but with these deviations we shall not be concerned. The justification of this procedure is the fact that for the analysis of price-making forces certain facts are not significant, because they are unlikely either to appear very often or to have very great effects when they do appear.

Market price is commonly contrasted with normal price. The latter is defined as the price which would prevail if only the more permanent forces affecting price-making were operative and if sufficient time were allowed to permit these forces to bring about a completely stable equilibrium of demand and supply. At this point we need not attempt an explanation of normal price, but the contrast between market and normal prices must be clearly grasped if the analysis of market price is to be understood. The analysis of market price differs from the analysis of normal price by taking into account somewhat different forces and by considering the short-run effects rather than the long-run effects of all forces. Cost of production is not, for example, a considerable factor in the determination of market price; but it is one of the most important in determining normal price, because as we shall see later, cost of production operates slowly, and in the short-time fluctuations of market price its influence is a minor factor. Fluctuations in the demands of buyers are exceedingly important in the analysis of market price but of less significance in the determination of normal price, because the passage of time allows producers to make adjustments in supply that largely — though not wholly — offset the changes in demand.

Demand and Demand Price. The simplest statement concerning the determination of market price is that demand and supply control it. This is the explanation most frequently heard in the market place and on the street. And the statement is true. For as the terms "demand" and "supply" are commonly used, they are taken to include all the forces that influence buyers and sellers. But like most other general and all-inclusive explanations this formula is of very little assistance in understanding the actual phenomena of price until the terms "demand" and "supply" have been exactly defined and until the complex forces for which they are but convenient names have been analyzed. Failure to observe these scientific procedures leads to many hasty and incorrect generalizations. For example, we are sometimes told that trade unions cannot permanently better the economic conditions of the working classes because "wages are determined by the immutable law of supply and demand"; or we read that because of the appearance of a monopoly "the law of supply and demand has been repealed"; or we are informed that "price having risen, the demand for the good fell off." Now as a matter of fact the principle of demand and supply does not preclude the successful establishment and operation of collective bargaining among laborers or of effective monopoly among producers. Neither is it true that demand falls off as a result of the rise in price. These and many other erroneous applications of the principle of demand and supply but illustrate the necessity for proceeding slowly and carefully in the analysis of the price-making process.

The first step in the analysis is the definition of "demand." "Individual demand" is defined as *the amount of any given good an individual stands ready to buy at a particular price, during a given time, and at a given place.* Demand is the amount the buyer *stands ready to take* or, more briefly, *will take,* and not the amount he has taken in the past. His demand cannot be exactly measured by computing the amount of the good sold him in the past, for this depends not on demand alone but also on supply. This distinction is of great importance. In the first place one may demand a good without making a purchase, for if the price he is willing to pay for the smallest obtainable quantity is below the figure at which the merchant is willing to sell, no sale can be made; yet his demand does unquestionably exist, and if the conditions of supply should change, the buyer who is now excluded might be able to buy. These excluded buyers are often of considerable importance. 1 A producer who can secure the economies of large-scale production must frequently rely upon them to increase the volume of sales when prices are reduced, and partly because of them the output of plants can be sufficiently enlarged to permit these economies to take place. Second, the amount 2 the buyer has purchased in the past is not an exact guide to the amount he will buy now, even though the price the sellers ask remains unchanged; his needs or desires may have altered, or his income may have become larger or smaller. Third, it 3 is logically incorrect to speak of demand as being the amount of a good purchased at a given price, since price and the volume of sales are the results of both demand and supply. The price the sellers set on their goods influences the volume of sales.

If demand is the amount of a good which the buyer stands ready to take at a given price, then the *individual-demand price* is the quantity of money which the buyer will give for a definite amount of a good. Price and amount mutually condition each other. The price must be stated before the amount the buyer stands ready to take has any significance, and the amount must be known before the price means anything. A housewife might be willing to pay a dollar a dozen for fresh eggs, but she would in all probability stand ready to take only a very limited quantity at that figure. On the other hand, she might be willing to buy two dozen a week, but she would not be likely to stand ready to pay a dollar a dozen for them.

The two constituent elements of demand are desire and ability to pay. Desire is fundamental, because where it does not exist goods cannot be sold at all. On the other hand, ability to pay is also essential, because without it demand cannot be made effective. When goods sell at low prices, the common conclusion is that these low prices are primarily due to the failure of the goods to awaken an intense desire on the part of the consumer. This conclusion is frequently correct, but it is not always so. Sometimes low prices result from the lack of purchasing power. Europe would gladly have used enormous quantities of American food products immediately after the close of the World War; but large numbers of her people were without ability to pay, and therefore their effective demand was relatively small. Some luxuries are demanded in quantities and at prices that seem entirely out of proportion to the objective usefulness inherent in them. Usually the explanation is that they are bought by the rich and the well-to-do, who are able to pay handsomely even for things for which they have no very intense desire. NB

The Individual-Demand Schedule. It is commonly observed that an individual buys less freely of any commodity when the price of that commodity is high than when it is low. Speaking generally, we find that when the price of any good decreases, the volume of purchases by the individual consumer increases and that it decreases when the price increases. This relationship between price and the volume of purchases, or more simply amount, gives rise to the individual-demand schedule, which is a list of the quantities of any given commodity the individual stands ready to buy at varying prices.

TABLE VII. INDIVIDUAL-DEMAND SCHEDULES

PRICE PER UNIT (CENTS)	AMOUNT BUYER WILL TAKE in given of	
	Commodity A	Commodity B
65		
60	100	
55	109	
50	120	
45	133	
40	150	100
35	171	102
30	200	105
25	240	110
20	300	120
15	400	150
10	600	300

Illustrative demand schedules of one buyer for two commodities are given in Table VII. The table shows that at 65 cents the buyer refuses to consider the purchase of either good; that at 60 cents he stands ready to buy sparingly of A, but that not until a price of 40 cents is reached will he take any of B; and that at relatively low prices he is prepared to buy both goods in much larger amounts. It must not be supposed that the average consumer is necessarily conscious of the different prices he would pay for different amounts of any commodity. If he has been accustomed to buy commodity A for about 15 cents, he is probably quite uncertain as to the amount of it he would be willing to buy at 50 cents. But the existence of a demand schedule does not require that he should be in possession of a plan of purchase for all probable prices. Studies of the marketing of goods show that even though consumers may not be conscious of the course of action they would take if the price of a commodity changed, purchases do decrease when the price of the commodity rises and increase when it falls.

Demand Curves. When two economic quantities vary in relation to each other, it is often convenient to illustrate their relationship by means of curves. In Fig. 4 the relationship between the prices and amounts of commodities A and B are thus illustrated. Two lines, *OX* and *OY*, are drawn to intersect at right angles at *O*. Price is conventionally measured by intervals on *OY* and amounts by intervals on *OX*. In plotting the two demand schedules given above the procedure is as follows: First, *OY* and *OX* are laid off at convenient intervals. (In this illustration price intervals of 10 cents and amount intervals of 100 units are convenient for the entire scale.) Beginning with commodity A, we find that 100 units is the smallest amount given in the schedule and that the corresponding price is 60 cents. So we insert a dot directly above the 100-mark on *OX* and directly opposite and to the right of the 60-mark on *OY*. Next, a dot is placed directly above the 200-mark and opposite the 30-mark, and so on until the final entry in the demand schedule is represented by a dot above the 600-mark on *OX* and opposite the 10-mark on *OY*. The next step is to draw a line connecting all the points, which is called the demand curve for commodity A. By similar procedure the points in the demand schedule for commodity B are determined and the curve representing that schedule drawn in. At 10 cents our hypothetical individual ceases to increase his purchases of A; and if the price is above 60 cents, he stops buying altogether. Hence we have terminated the curve A at these points.

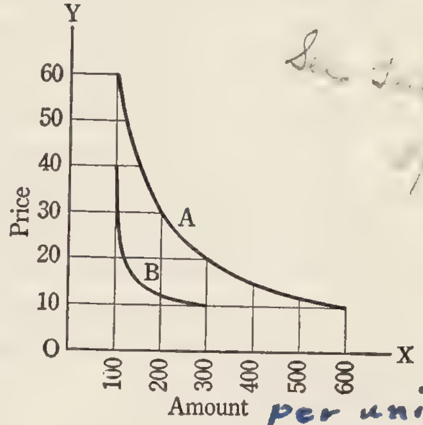

FIG. 4

See [illegible]

per unit of time

The advantages of translating demand schedules into demand curves, as well as certain dangers in the procedure, must be observed at this point. From the schedule it is seen that when the price falls from 20 cents to 15 cents, the amount of A that will be taken increases at the same time from 300 units to 400 units. As far as the schedule shows, we know nothing about the amounts that would be taken at prices between 20 cents and 15 cents; but when we draw a curve, we do actually represent the amounts which would probably correspond to the smaller price intervals. We assume that if the amount increased at a certain rate when price fell from 20 cents to 15 cents, its increase between these price points was regular. Since the amount interval in the schedule corresponding to the price interval of 5 cents is 100, we assume that for every decline of 1 cent in price the amount increased about 20 units. Where the entries in the demand schedule — either price or amount — are far apart, the curve enables us to gain some notion as to the amounts that would be taken at intervening price points.

Suppose we wish to find the amount of A that the buyer would take at $17\frac{1}{2}$ cents. The first step is to ascertain the point on the demand curve, *A*, which corresponds to the price of $17\frac{1}{2}$ cents on *OY*. Then the point on *OX* directly below this point is ascertained. In this case we find that a price of $17\frac{1}{2}$ cents corresponds approximately to an amount of 350 on the axis *OX*. We therefore infer that although the schedule does not show an entry of any amount opposite $17\frac{1}{2}$, the buyer would probably take 350 units at that price. In coming to this conclusion we have, however, assumed that if the price declined from 20 cents by 1-cent intervals, amounts would increase by intervals of 20 units. As far as the schedule of demand for A is concerned, this may or may not be true. For large numbers of consumers' goods, such as common kinds of food, for many minor items of clothing, and for a multitude of small articles of consumption, our assumption of continuous and regular variations of amounts between the two price points is essentially correct. For these goods the curve represents the actual conditions with a considerable degree of accuracy. On the other hand, it is well established that the curve does not represent the facts in respect to articles which are purchased only occasionally and in relatively large, costly units. A man may buy a pair of shoes each season of the year at $10 a pair and decline to buy an additional pair at any price above $5 or $6. Or a family may install a telephone in their dwelling at a rental of $5 a month and refuse to have a second put in even though the company changed its rates and offered to install and operate at $2 a month. Of course it might well happen that if the household were conducted on a more elaborate scale, a second connection would be taken at only a small reduction in price. To sum up: For some commodities a curve does not represent the demand schedule of the individual, because the variation in amounts he stands ready to take is not continuous but moves by large jumps instead of creeping by small intervals. The correct representation for such a demand schedule is a series of crosses or dots to show the variation. For other commodities there is continuity, and small price changes will be accompanied by small variations in amount.

Collective Demand. Among the consumers who patronize the same market there are many who demand the same good. Taken together, they make up the collective demand for it, and the amounts which they stand ready to buy at different prices make up the collective demand schedule. There is of course much similarity between the demand schedules of a large number of individuals buying the same commodity in the same market, yet there is always some degree of dissimilarity. There are differences in tastes, family requirements (because of differences in the size, age-composition, and standards of consumption of the different families), family incomes, and occupations. Hence variations in the collective-demand schedule are not similar to variations in the demand schedule of any individual except by chance. Thus, in the preceding example we found that the individual would increase his buying by 33 per cent if price fell from 20 cents to 15 cents, but it is improbable that the same change in price would affect the buying of the entire consuming public in exactly the same proportions. This lack of similarity between a collective demand schedule and any of its component individual schedules is illustrated by Table VIII, in which three individual-demand schedules are summated to make a collective-demand schedule. These schedules are represented graphically in Fig. 5 on page 109.

TABLE VIII. INDIVIDUAL-DEMAND AND COLLECTIVE-DEMAND SCHEDULES

PRICE PER UNIT (CENTS)	SCHEDULE I (AMOUNT)	SCHEDULE II (AMOUNT)	SCHEDULE III (AMOUNT)	COLLECTIVE SCHEDULE IV (AMOUNT)
60	80	20		100
50	90	30		120
40	100	40		140
30	110	50	10	170
20	130	70	20	220
10	160	100	50	310

Characteristics of Demand Schedules. If we investigate the amounts of different commodities sold within any market, we find that there is the greatest variation. Some articles are sold in very large amounts, and others are purchased only in very limited quantities. Some can be sold only if their prices are relatively low, and others, although they are taken in larger amounts when their prices fall, can be disposed of in large quantities even when their prices are relatively high. From these facts we conclude that the demand schedules of the individuals who make up the entire body of buyers are unlike in respect to different goods. Referring again to Fig. 4, p. 105, and the accompanying demand schedules, we find that curve *A* differs from curve *B* in several important ways. It has greater extent, that is, it shows that the buyer will take the good through a greater range of prices; and it has greater volume than the other curve, that is, it shows that the buyer will take larger amounts at given prices.

The demand schedule represented by curve *A* also has greater elasticity at some prices and less at others than that represented by curve *B*. The demand for a commodity is said to be *elastic* when a relatively small change in price is accompanied by a relatively large change in the amount the buyer (or all buyers collectively) stands ready to take. If, on the other hand, the change in price is accompanied by a relatively small change in the amount the buyer stands ready to take, the demand is said to be *inelastic*. For example, if a drop of 1 per cent in price is followed by a much larger percentage increase in the volume of sales, the demand is said to be very elastic; but if a 1 per cent decrease is followed by an increase in volume of sales of much less than 1 per cent, the demand is very inelastic. If the rate of increase in sales exactly equals the rate of change in price, the demand is said to have an *elasticity of one*. In terms of Fig. 4, the demand for commodity A has an elasticity of one throughout its entire extent, and the demand for B has an elasticity of less than one between 40 cents and 15 cents, and an elasticity of more than one between 15 cents and 10 cents.

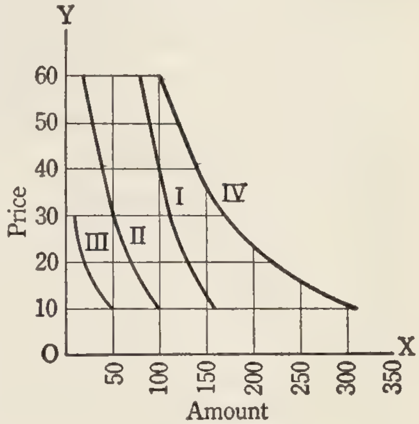

FIG. 5

The elasticity of demand cannot be tested by determining whether the *absolute* increase in sales is great when price falls a certain amount or by determining whether the decrease is absolutely large when price rises. In Fig. 5 the demand represented by Schedule II is relatively elastic between the prices 60 cents and 40 cents, although the absolute increase in amount demanded is only 20 units. The demand represented by Schedule I is *relatively* inelastic between these prices in spite of the fact that the *absolute* increase is exactly the same as in Schedule II. In Fig. 4 the elasticity of demand for commodity A is about the same between the prices 15 cents and 10 cents as it is between 60 cents and 55 cents, although the absolute increase in amount demanded is 200 units for the former price interval and 10 units for the latter. The change from 60 cents to 55 cents is *relatively* small, and it only needs to be accompanied by a *relatively* small change in amount to conform to the definition of elasticity of one. On the other hand, the change in price from 15 cents to 10 cents is 50 per cent of the lower price and must be accompanied by a proportionately large increase in the amount to agree with the definition.

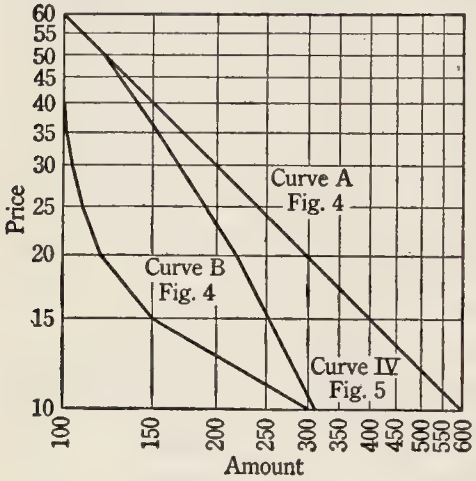

FIG. 6

It is evident that one can judge the elasticity of demand from curves plotted on an absolute scale, as in Figs. 4 and 5, only with the greatest difficulty, because an ordinary arithmetic scale does not show *proportional change*. In Fig. 6 the two demand schedules, which are shown on absolute scales in Fig. 4, have been plotted on logarithmic scales, which are arranged to show proportional variations of both price and amount; that is, the intervals on the amount scale (corresponding to OX in Fig. 4) are so spaced that the distance on the chart between 100 and 200 is exactly the same as between 200 and 400 (since 100 is to 200 as 200 is to 400). On the absolute scale the interval between 200 and 400 is twice as great as between 100 and 200. The intervals on the price scale in Fig. 6 (corresponding to OY in Fig. 4) are likewise so constructed that a given linear distance represents the same proportional change; for example, the distance between 10 cents and 20 cents is the same as between 20 cents and 40 cents.

When the individual-demand schedule for commodity A is plotted in Fig. 6, it is seen to be a diagonal straight line. With every variation in price the change in the corresponding amount is exactly proportional. As price falls from 40 cents to 15 cents, the amount of B increases less than proportionally, as is shown by the fact that the curve falls faster than it moves to the right. But from 15 cents to 10 cents it moves to the right faster than it moves downward.

Conditions which affect the elasticity of individual demand Imp
are as follows: First, the individual demand is usually elastic 1
for luxuries and inelastic for necessities. We tend to buy necessities near to the point of satiation before we branch out on expenditures for conveniences and pure luxuries. In order to maintain life and the decencies of civilization we must have a certain amount of ordinary food and clothing. Few of us would desire to consume very much more of these essentials even though their prices declined a great deal; nor would we diminish our consumption rapidly if prices rose. Second, 2
the individual demand for a good is likely to be elastic if other goods can be substituted for it or if it can be substituted for other goods. There is scarcely any substitute for white bread among a majority of the population in the United States, and accordingly the individual demand for it is relatively inelastic; but the demand for a single brand of breakfast food is elastic by comparison because other brands can take its place just as it can take their places in the consumption of most individuals. A large change in the price of one brand is likely to be followed by a greater change in the volume of purchases than would follow from a change in the price of salt, brooms, telephones, and hundreds of other articles or services of daily use for which substitution is difficult. In 3
the third place, the demand for a good tends to be more elastic if it has a variety of uses than if it can serve us in only one capacity. This statement is more generally true for raw materials than it is for finished goods because, after all, the latter are not as a rule adaptable to a large variety of uses.

But a good such as coal has a multitude of specific uses in industry in the manufacture of a vast variety of commodities. Fourth, the demand for any commodity is likely to be
4 inelastic at very high and at very low prices. When it is costly, people persist in using small amounts; and when it is cheap, they buy it to the point where the desire for it is nearly
5 satiated. Fifth, demand is likely to be inelastic when consumption is greatly influenced by habit and custom.

The elasticity of the collective-demand schedule is determined by the same conditions that affect the elasticity of its component individual schedules. But it cannot agree with them all. It may be more elastic at a given price than some and less elastic than others. If new buyers enter the market in large numbers at a certain price, the elasticity of the collective schedule may be greater than that of any of its components. The elasticity of an individual-demand schedule cannot be important, since even a small dealer supplies a number of buyers. It is the elasticity of the collective-demand schedule with which he is concerned. Upon the degree of elasticity of collective demand depend a number of important business decisions. Would it be profitable to cut prices in order to expand and take advantage of the economies of larger production? If the demand is elastic, the prospects of securing enough additional sales to enable the producer to reduce costs are very good; but if it is inelastic, they are poor. Or suppose that the producer's costs are rising. Should he prepare to reduce output greatly as prices are forced up by rising costs? If demand is inelastic, he need not reduce greatly; but if it is elastic, he may be forced to shut down a large part of his plant.

Changes in Demand. Since demand has been defined as the amount of a certain commodity an individual or a group of individuals stand ready to take at a given price during a given time and in a given market, an increase in demand must mean that the buyers will take a larger amount at a given price; and a decrease in demand, a smaller amount. When the attitude of the buyers toward the good so changes that they will take more or less at one price in their schedule, it usually happens that a considerable part of the demand schedule and demand curve will be altered. Thus, if we assume that the demand for a good as represented by the collective-demand schedule on page 108 has changed from 140 units at 40 cents to 160 units, we should expect to find that more would also be taken at 30 cents and 20 cents, as well as at 40 cents. When it is said that the demand for a good has changed, it is implied that a considerable portion of the demand schedule has moved. But it must not be assumed that all parts of the schedule and of the corresponding curve have changed in the same proportion. In Fig. 7 the original position of a collective-demand schedule is represented by the curve AA. The curve $A'A'$ shows an increase of demand throughout the entire schedule, and $A''A''$ shows a decrease. Such modifications as CA and AD represent partial shifts in the original schedule.

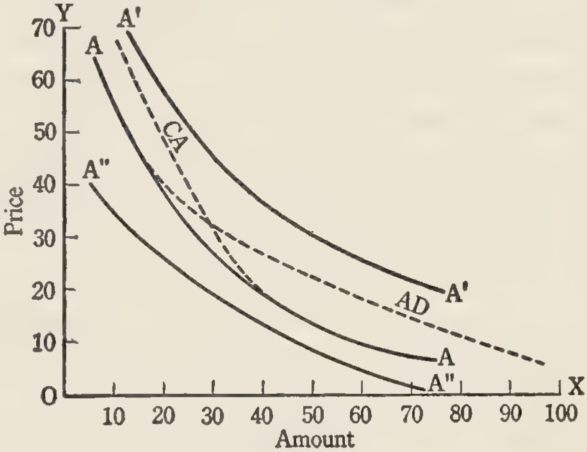

FIG. 7

A common misconception of the meaning of "an increase in demand" is illustrated in the statement "Because of a decline in price, the demand for automobiles has considerably increased." If the statement is taken to mean what it says, we should infer that price determined demand and that dealers were entirely independent in setting prices. But this is obviously not true. The true relationship between price and demand, which is understood by a majority despite the prevalence of such ready phrases as the one just quoted, is reciprocal.

If we look at Schedule IV in Table VIII on page 108 we can see that if the dealers reduce the price from 40 cents to 20 cents, the amount the buyers will take increases from 140 units to 220 units without any change in demand.

An increase in the volume of total sales of any commodity in any market does not always show that an increase of demand has taken place. If the sellers have reduced the price of the commodity, we should infer that the additional sales were brought about by increased purchases by old buyers and by the entrance of new buyers who were unwilling to take the commodity at the old price. The true test of whether demand has actually increased can be made by putting price back to the old figure. If the volume of sales is found to be greater than before, we shall then conclude that demand has increased. But when sales grow with the reduction of price, we cannot be sure that the additional purchases are not the result of a latent demand which existed before the sellers decided to cut the price.

Supply and Supply Schedules. In preceding sections we have stressed the point that demand does not mean merely amount purchased and that it should not be taken to mean readiness to purchase a given amount unless the price at which potential buyers are willing to take that amount is also stated. The same points must be borne in mind in analyzing the concept of supply. There are, it is true, instances in which stocks of goods are thrown upon the market to be sold at auction to the highest bidder. But in these instances the price at which the goods are offered before the sale opens is any price above zero. Actually there are occasions on which goods are so offered. But usually the sellers have one or more other opportunities to dispose of their stocks if the auction does not call out buyers at higher prices. Hence we conclude that, as in the case of demand, supply always implies a relationship between amount and price. "Supply" is therefore defined as *the amount one seller or a group of sellers stand ready to offer during a given time, in a given market, at a given price.* The amount one seller will offer is an individual supply, and the amount all the sellers taken together will offer is a collective supply.

The price at which any seller stands ready at any time to offer a given amount of any good is his *supply price* for that amount. And the price at which all the sellers in the market stand ready to offer a given amount is the market, or collective, supply price. The individual-supply schedule is a list of the varying amounts an individual stands ready to offer for sale at different prices. The collective-supply schedule for any commodity is made up, like the collective-demand schedule, of the individual schedules. When supply schedules are represented by supply curves, it is found that instead of slanting downward from left to right the curves slant upward from left to right; that is, as price rises, the tendency is for offers to become more numerous and for more and not fewer goods to be put upon the market.

Supply schedules and curves for different sellers of the same good are usually somewhat unlike, just as the demand schedules of different buyers are unlike. Hence the collective-supply schedule is a composite of unlike and divergent individual schedules. The supply schedules, both individual and collective, for different goods are also unlike; that is, a given increase or decrease in the prices of several goods will cause the amounts offered to increase or decline in different proportions. The differences between individual schedules are illustrated in Table IX, in which A, B, and C are amounts of the same good offered by different sellers in the same market.

TABLE IX. INDIVIDUAL-SUPPLY SCHEDULES

PRICE	SUPPLY A	SUPPLY B	SUPPLY C
10	2000	2000	5000
20	2500	2000	5000
30	3000	2000	5000
40	3500	2250	5000
50	4000	2500	5000
60	4500	2500	5000
70	4500	2500	5000

In the case of highly perishable goods already in the consumers' market the sellers can scarcely be said to have supply schedules. Any amount of the good they happen to have on hand they must either sell or allow to spoil. If competition is perfect, they will sell the entire stock for whatever price it will bring. This conclusion appears to be contrary to fact, since it often happens that dealers in perishable goods maintain their established prices and sell whatever *quantity* they can, allowing the remainder to spoil. And it is argued that this course is pursued by the dealers because (1) it is more profitable for them to sell a smaller amount at a higher price than they could obtain for their entire stock and (2) because they are afraid of "spoiling the market." The first reason requires for its validity an assumption that competition is far from perfect. Wholesale dealers in fresh fruits might indeed find it more profitable to sell three fourths of their supply at a price of five dollars a unit than the entire stock at three dollars a unit. But if every wholesale dealer in a large market had equal access to every retailer and if every retailer had immediate access to the consumers of several of his competitors, no dealer could sell any of his supply at five dollars; for so long as there existed any unsold stock in the hands of any dealer, he would try to take business away from his competitors by lowering the price a trifle. But the competitors would reply by lowering theirs a bit more than his, and the final result would be the fixing of a wholesale price at three dollars. Why, then, do we find that dealers maintain prices and allow goods to spoil? The answer is that each wholesaler has a line of retailers to whom he sells regularly, and each retailer has a group of consumers who customarily patronize his store. If the wholesalers find that they have overbought, they tend to sell what they can at the established price or a little below, because that policy yields them a higher gross return and a higher profit, as will be explained presently. They know that other wholesalers cannot quickly invade their line of retailers and take away the business. The retailer does not have established relations with many wholesalers, and he must buy perishable commodities almost daily and in small amounts. For both reasons he tends to take these commodities from the wholesaler at the latter's price, regulating the amount he takes according to the amount he thinks he can sell at the wholesale price plus his customary margin.

But this condition of static price cannot continue indefinitely. The growers will bring the commodities to the wholesale market as long as they can obtain a price which will cover costs of harvesting and transportation. The wholesalers will presently lower their buying price, and the consumer will buy at a price somewhere near the price which will clear the market. The one condition that will prevent the tendency toward the market-clearing price is the existence of a monopolistic agreement among the growers. If this occurs, it is entirely possible that considerable quantities of perishable foodstuffs will annually be allowed to go to waste. But here again the quantity will not continue to be very large, because the growers will limit their production to avoid the cost of producing the amount which is wasted.

The phrase "spoiling the market" means that the seller may change the consumers' habits of consumption and their buying behavior to his disadvantage. For some commodities the consumer comes to have in mind a price at which he will buy freely, another price at which he will buy sparingly, and still another at which he will not buy at all. And these prices are much affected by his experience in the past. If the consumer has been accustomed to buy a fresh fruit in season at from 15 cents to 20 cents a quart, and if for a period of a week the price drops to 10 cents, the dealers may have difficulty in selling to the same consumers the following week at 15 cents or 20 cents. Here we encounter a very complex problem. Account must be taken of the income of the typical consumer, the number of substitutes for the particular good in question, the stability of price in the past, — that is, whether the consumer has been accustomed to pay rather stable prices for this good or has been accustomed to pay varying prices, — the extent of competition among retailers, and many other facts.

Some goods which do not deteriorate physically are subject to rapid deterioration of utility with the passage of time. Such "style goods" as articles of dress are the stock instances. A retailer who has purchased a stock of goods of this character must sell it before the style changes or the season in which it is wanted has passed. And here again the supply is fixed to a great extent, and the seller must take what he can get. Likewise the dealer will try to protect himself by maintaining prices and by allowing some of the commodity to be held past its season.

There is of course a tendency for manufacturers and dealers of all sorts to place themselves in a position where they can either withhold goods from the market if the prices are deemed too low or augment stocks if prices are found to be higher than expected. Refrigeration helps to accomplish these ends for such goods as perishable foods, and larger storage facilities at central marketing points do the same for nonperishable goods of all sorts. For some kinds of manufactured commodities surplus factory capacity — that is, a surplus relative to normal requirements — serves the same ends.

For relatively durable goods not affected by changes in style, supply schedules of some range do exist. A farmer must of course sell his crops sometime during the year following the harvest. But he may distribute his sales within this time according to his judgment of the profitableness of selling at one time rather than another and according to the cost and convenience of marketing crops at one time rather than another. Still another consideration that may affect his decision to sell is his need for funds to pay outstanding debts incurred in producing the crops. Hence he may sell a part of his wheat direct from the thresher, reserving the rest until he thinks the price is the best that he is likely to get.

In case the goods are durable finished commodities or half-manufactured goods, supply schedules exist; but they are doubtless much restricted under competitive conditions. For although the goods are not perishable in a physical sense, producers and dealers do not wish to hold them very long in storage because by so doing they incur the risk of a decrease in price and because storage and interest charges accumulate while the goods are held in stock. Hence when dealers are confronted with a drop in the prices of particular wares, they ordinarily sell the stocks they have on hand for whatever prices they will bring. It is also to be observed that a sudden unanticipated increase in the demand for a given commodity will not always quickly call forth a larger volume of goods. The flow of commodities into the market is regulated in advance by the producers, and time is required to make additions to the rate at which they are fed into the market.

There are certain goods and services for which the supply schedules are very extensive; that is, they run through a wide range of prices and amounts. Dwelling houses, used automobiles, furniture, and other goods which are frequently offered for sale by consumers who have the alternative of selling or using them are commonly offered at greatly varying prices. Because different people are certain to attach to these goods different values for personal use, there is less uniformity than in a market where only dealers are sellers. The supply schedule for labor is also extensive. Here the wage rate is weighed against the irksomeness of labor. At low wages (low relative to the established rate) many laborers refuse to work at all and others refuse to work as many days in the month, whereas higher wages call forth more workers and induce many men to work overtime and more days in the month.

In Fig. 8 on page 120 are plotted three schedules of supply shown in Table IX on page 115. The perpendicular lines extending from the termini of curves *A* and *B* serve only to show that after certain prices have been reached, amount ceases to vary with price. In the case of supply C, the amount is fixed and there is no covariation and no curve, but simply an ordinate; that is, a perpendicular line extending indefinitely above the axis *OX*.

The Determination of Supply. In our discussion of the types of supply schedules, we noted that sometimes a stock of goods is thrown upon the market to bring whatever the buyers will give for it. This is sometimes called supply without reservation, because no part of the stock is withheld if the price is unprofitably low. In many other cases sellers do reserve goods; that is, they hold them at certain prices below which they refuse to sell. The conditions which govern the fixing of dealers' supply prices are very complex, and we cannot go into them at length at this point. Briefly stated, the following are the most important: (1) dealers' estimates of the prices consumers will probably pay for the amounts available now and at a near future date; (2) dealers' estimates of the amounts of the good that will come upon the market in the near future; (3) the prices of substitute goods; (4) present and probable future costs of selling and storage; (5) the prices at which competitors now offer, or are likely to offer, the good in the near future; (6) present and future estimated cost of buying goods from other dealers or producers.

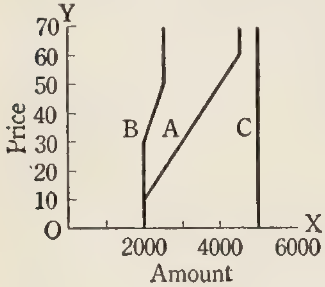

FIG. 8

The seller in a modern competitive market must be prepared to change the reservation prices he has set upon his stocks; for such reservation prices are largely estimates on the part of the seller of the prices the market will pay for the goods. But no one can predict these prices accurately in advance of actual experience. Neither the exact demand prices of the buyers nor the number of people who will buy can be forecast, and the amounts of goods in the hands of manufacturers, competing dealers, and other marketing agencies cannot be easily or accurately ascertained. The supplies of substitutes and the degree to which substitution can take place are also indeterminate. The practical method of pricing seems to be that of "cut-and-try." If the price which is first tried is too high, few goods will move; and that is sufficient warning to the sellers to reduce prices. It often happens that they must reduce them so low that the goods they have stocked become the occasions of loss. If this is the case, smaller quantities will be stocked in the future; and in the long run the goods will either sell for enough to enable manufacturers and dealers to make and handle them profitably or they will disappear from the market.

The difficulty of ascertaining in advance the proper reservation prices causes the supply schedules of many goods to be extremely variable. There may be no such thing as a stable price. In other lines, particularly in the sale of branded goods and small wares, reservation prices change only infrequently, the manufacturer preferring to regulate the supply so as to feed the market at a rate which will maintain the established price. The reservation prices attached to goods or services offered for sale by monopolies are also likely to be stable.

Changes in Supply. Increase in supply means that more goods are offered at a given price than before the increase took place, whereas a decrease in supply means that fewer goods are offered at the given price. An increase or a decrease of supply at one price in the supply schedule usually means that the supply has increased or decreased at adjacent prices; and it is true for supply, just as for demand, that a change in the schedule at a given price point usually involves changes for other near price points. Hence an increase of supply is illustrated graphically by placing the supply curve farther to the right and below its original position, a decrease by placing it farther to the left and above its original position. But it must be remembered that because the supply curve (or schedule) has changed for certain prices, it does not always follow that it has changed for all prices.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Can the price of a good change without a change in its value? How?

2. It is often said that a good cannot have two prices in the same market at the same time. Under which conditions is this statement correct? Under which, if any, is it not correct?

3. The sales of automobiles have increased greatly since 1917. What facts can be given to show that this increase is not entirely owing to an increase in demand?

4. State whether you should expect the collective demand for the following goods to be elastic or inelastic: (1) gasoline, (2) cotton cloth, (3) tobacco, (4) a popular novel. Give reasons.

5. What is meant by "the continuity of the demand curve"? List several articles for which the individual demand is not continuous; for which it is continuous.

6. Are rapid changes in demand desirable or undesirable from the standpoint of producers?

7. Are changes in supply desirable or undesirable from the standpoint of the consumers, if it is assumed that the average supply price for a commodity remains the same?

CHAPTER VIII

THE PRICE-MAKING PROCESS: MARKET PRICE

Reserved and Unreserved Supply. The supply of a good is *reserved* if the seller offers it for sale at a predetermined price. And the price which the seller sets upon such a reserved supply is called the *reservation price.* In order that sellers may reserve their supplies and set reservation prices upon them, they must be able to hold back their stocks if prices are in their judgment too low, and to push them forward when prices are satisfactory. Reservation implies, therefore, the existence of storage facilities and waiting power on the part of the sellers.

The supply of a good is *unreserved* if the sellers throw it upon the market without setting any price whatever upon it. Often reservation prices are not made public. The sellers simply ask the prospective buyers, "What will you give for this commodity?" But in a majority of these cases those who offer the commodity have privately determined that if the price does not meet their expectations, they will reserve their stocks either until the price rises to what they are willing to take or until their reservation price has been proved too high.

Price without Reservation of Supply. The determination of price without reservation is the simplest price phenomenon encountered in the market, and in many markets it is nearer the actual conditions than is the assumption of reservation of supply. Hence the making of price without reservation will be explained before the more complicated problem of price with reservation.

The conditions of price-making without reservation are as follows: (1) the demand for the good is given, and (2) a given quantity of the good whose price is being studied is thrown on the market to sell for whatever it will fetch, any price above zero being acceptable to the sellers. The first effect of throwing an unreserved supply into the market depends upon whether the buyers have more or less complete information of the quantity being offered. If they believe that the quantity is small, they may pay high prices for a time and, when the amount of the good becomes known, lower their bids for it. It may also happen that they have overestimated the supply and that the initial low price will be followed by higher prices later on. But when this preliminary higgling has established the actual amounts for sale, price will settle down to a more stable level. It is with the stable level that we are here concerned.

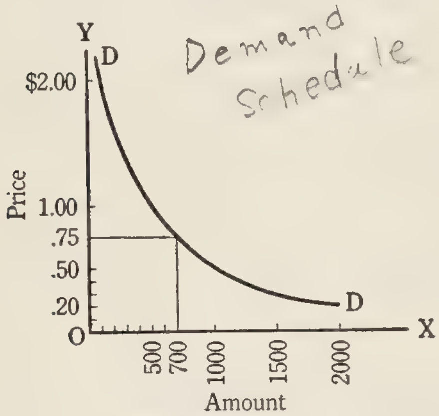

FIG. 9

If the demand for a good is given and if a definite quantity is offered for whatever it will bring, then the level toward which the higgling will tend to drive price is determined by the collective-demand schedule for that market and by the amount offered.

In Fig. 9 the demand schedule, which is represented by *DD*, varies from 100 units at $2 to 2000 units at 20 cents. If, now, the sellers throw 500 units on the market without setting a reservation price, the good will sell for $1 a unit, and the gross receipts from that amount will be $500. If 1000 units are offered, the price will be 50 cents a unit and the gross receipts $500, or exactly the same as though 500 units had been sold. And if 2000 units are offered, the price will be 20 cents and the gross only $400, or $100 less than if one fourth that quantity had been marketed.

It is evident that when the demand schedule is fixed, —

that is, does not move to the right or the left in terms of the diagram, — any increase in the amount offered will lower the price, and any decrease will raise the price. These conclusions follow from the fact that the demand curve always slants outward from *OY* and downward toward *OX*. In the illustration it appears that a large amount not only brings a smaller price per unit, but also a smaller gross return. But this is not always the case. In the first place, the smallest amount that could be sold with demand as represented in Fig. 9 — 100 units — will sell for $2 a unit and yield a gross return of only $200, which is less than if 500 units had been offered. A more careful inspection of the prices at which various amounts can be sold shows that the maximum gross is received by the sellers when approximately 700 units are sold for a price of 75 cents, which will return $525 gross. Any amount in excess of 700 units yields less in total money return, and any amount below 700 will do the same. The reason for this condition is plain. Gross receipts are the product of price and amount, and at either end of the demand curve one of the two factors — price or amount — approaches zero, and their product also becomes zero. At the upper end of the curve, price is very high, but only a negligible quantity can be sold; at the lower end, large amounts can be sold, but the price becomes very low. Hence it follows that between the two extremes, at some mid-point in both price and amount, the gross receipts will be at a maximum.

This conclusion applies to all goods sold in markets where the demand schedule has definite limits. If any good were an absolute necessity of life, its demand schedule would show that indefinitely high prices would be paid for small amounts, since life itself would depend upon it. The demand curve, instead of meeting *OY*, would become nearly parallel to it and extend upward indefinitely. Under this condition a relatively small amount would bring a fabulously high price and yield the maximum gross receipts. This condition of demand is shown by the curve *EE* in Fig. 10.

The opposite situation to the one just described would be encountered if it was assumed that at low prices the demand curve flattened out and became parallel to *OX*. Curve *DD* in Fig. 10 shows that the larger the amount put upon the market the larger will be the gross receipts. Even though the price falls with each increase in sales, it falls slowly while amount increases rapidly. Of course there is no commodity whose demand is so elastic that after a certain point has been reached amounts taken increase very greatly and indefinitely with each minute decline in price. All demand schedules reach zero price much short of the point where more than a small proportion of the nation's productive energy can be devoted to any single good.

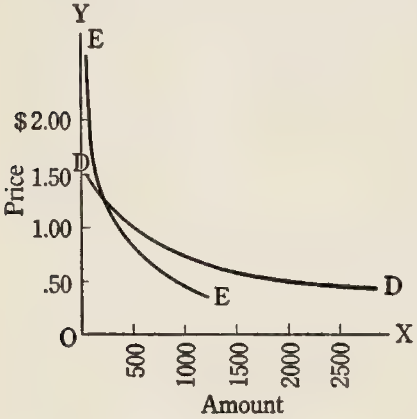

FIG. 10

But it is not contrary to fact to say that at certain points in the demand schedules for some commodities the amounts buyers stand ready to take increase so rapidly with small changes in price that gross receipts also increase rapidly. If the demand for a good is sufficiently elastic between two price points, gross receipts must increase. Such a condition of elasticity depends on the possibility of direct substitution of the good for one or more other goods by a large number of people. But when this condition applies to a certain good, it must be true not only that the good is readily accepted by the consumer in place of other goods but also that the producers of the other goods cannot reduce their prices enough to prevent the substitution.

The effect of high elasticity upon price and gross receipts must be sharply distinguished in this analysis from the effect of increasing demand. In Fig. 11 the demand schedule moves from an original position *DD* to *D′D′* and *D″D″*, the latter two positions both representing increases in demand at all points on the demand schedule. When the demand is at *DD*, 500 units sell for $1 and gross receipts are $500; when it is at *D'D'*, 1000 units sell for $1 and gross is $1000; and when it is at *D''D''*, 1500 are sold and gross is $1500. The conditions represented in Fig. 11 are quite different from those illustrated in Figs. 9 and 10. In the latter two figures the demand was assumed to remain constant while the amount offered increased or decreased. In Fig. 11 both the demand and the amount offered have changed, and as a result of this double change the price has remained the same. But the stability of price is due solely to the fact that both the demand and the amount offered have changed. If the amount had lagged behind demand, price would have risen; if it had outstripped demand, price would have fallen.

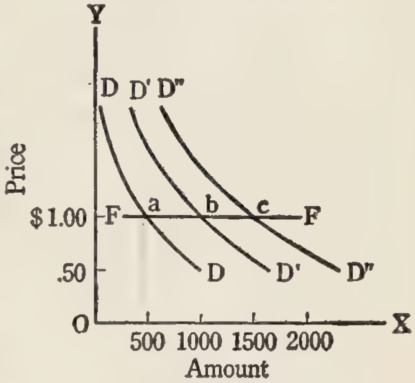

FIG. 11

Such increases in demand as are presented in Fig. 11 result from the natural growth of population, from immigration, from an increase in the national income, or from changes in habits of consumption which involve the abandonment of some commodities and the adoption of others. But these are all dynamic changes which do not happen within a short range of time. In a short period of time the sellers do not have the choice of selling either 500 or 1000 units at $1 a unit. Having produced either quantity and being unable to reserve supply, they must sell for the price which is determined by the collective-demand schedule.

Thus far in the discussion we have assumed that goods are thrown upon the market without reservation — without any price set upon them by the sellers. It may well be objected that sellers practically never do this. Although the sellers usually set up, either openly or in their own minds, the minimum

prices they are ready to accept for different amounts of the goods they sell, several conditions actually encountered in economic life detract somewhat from the significance of reservation prices. In the first place, the sellers may be unable
1 to forecast the demand. They may expect that it will be sufficient to carry off 1000 units at $1 and produce $1000 gross; but when this amount is brought to market, it may be found that instead of being at *D′D′* (Fig. 11) demand is at *DD* and the stock which has been produced must be sold for a gross of $500. If this is the case, it is futile for the sellers to hold the good at $1, provided it is necessary to dispose of the entire stock within the period of time to which the demand curve *DD* is applicable and provided competition exists. But if the good is durable and if future production can be controlled, the sellers may find it more profitable to offer something between 500 and 1000 units and hold the remainder for a higher price in the future. If the future supply cannot be controlled or if
2 the stock is perishable, they must take 50 cents. To be sure, reservation prices will be attached to the good in either case, but in the latter they are derived entirely from the demand as shown by the prices at which consumers are willing to buy.
3 For example, farmers cannot store more than a small portion of any crop. If in any given year the wheat crop happens to be unusually large, it must be sold for the price which is determined by the world demand for wheat and by the amount of the harvests in all the wheat-producing areas of the world. A small amount can be held in storage; but since next year's planting cannot be reduced materially and since next year's yields, which depend largely on the weather, cannot be effectively controlled, this procedure is not of material benefit.

The danger from excessive production is not confined to agriculture. Within a short period of time the market for many commodities is limited. Demands do not have a high degree of variability, because, among other reasons, consumption is so greatly controlled by habit and custom. Hence the principles of price which apply to supply without reservation are significant for price analysis in all types of production. The sellers' supply curves for short periods of time are short and are largely the results of variations in the estimates of sellers as to the rate at which goods are flowing to the market and the rate at which buyers are willing to consume.

The Significance of Supply Prices. The system of uniform prices for all buyers, which has now superseded the older system of individual bargaining throughout retail trade in the United States, is essentially a device for economy and for retaining goodwill. It eliminates time formerly wasted in 1/2 making individual bargains; and it pleases the buyer, who likes to believe that he is receiving equal treatment with all the other patrons of the store. The buyer must pay the merchant's price or go without the good. This system gives the appearance of greater power to control price than the seller actually possesses, but it must be remembered that a seller cannot maintain a predetermined price unless he can also limit the amount the entire market offers for sale. If the consumers do not want to pay the price asked, they can easily force the seller to reduce it by refusing to buy.

The actual control over prices, which the seller exercises N, B quite as much as the buyer, arises from the control of quantities and not from the power to set reservation prices. Only when the seller has correctly adjusted his offered amounts to the prices he asks, can these prices be maintained. Hence, if any producer wishes to control price, he must limit the amount of goods offered for sale. Practically all consumers' goods and many producers' goods are fed into the market at rates which the sellers believe will maintain certain prices. But the maintenance of these prices frequently fails, because it is often impossible for manufacturers and merchants to control the total amounts offered for sale and because the demand is very often incorrectly forecast.

Market Price with Reservation of Supply. In the preceding sections it was shown that the establishment of reservation prices results in the supply schedule and the supply curve.

The supply schedule represents the amounts that sellers stand ready to sell at varying prices, and the demand schedule the amounts that buyers stand ready to buy at the same prices. When amounts offered for sale in any market are regulated by reservation of supply, the price at which the good sells depends upon the relationship of the demand schedule and the supply schedule. In Table X the schedules of demand and supply show how market price is determined under these conditions.

TABLE X

PRICE (CENTS PER UNIT)	DEMAND SCHEDULE (UNITS)	SUPPLY SCHEDULE (UNITS)
11	700	2300
10	900	2150
9	1100	2000
8	1400	1850
7	1700	1700
6	2000	1450
5	2600	1250
4	4000	950

At the relatively high price of 11 cents the rate of flow of goods to the market is more rapid than the rate at which consumers are willing to buy, and at the relatively low price of 5 cents the amount the consumers stand ready to take far exceeds the amount the sellers care to offer. But at the price of 7 cents the amount sellers offer and the amount consumers stand ready to buy are exactly equal. This is the only price which will bring about an equilibrium between demand and supply. There is, then, in any given market, under competitive conditions, a single price at which demand and supply are equal; if any other price happens to be set at any time, it cannot be maintained. Price will move to the point where demand and supply are equal.

Effects of Changes in Demand; Supply Constant. In Fig. 12 the demand and supply schedules of Table X are represented by the curves *DD* and *SS*. The equilibrium price is seen to be 7 cents, or *NP*, and the amount sold, 1700 units, or *ON* units.

When demand increases throughout the entire schedule, the new equilibrium (the supply schedule remaining the same) is at the intersection of *SS* with *D'D'*, or at *Q*. The new price is *MQ*, or 9 cents a unit, and the amount sold is *OM*, or about 2000 units. This result is different from the effect of an increase of demand when supply is a fixed stock. In the present illustration an increase of demand with only 1700 units for sale would have sent the price to about 10 cents a unit. But since the sellers were willing to offer 2150 units for 10 cents a unit before the rise in demand took place, the price would not rise to 10 cents but only to 9 cents. It can readily be seen from the diagram that if demand declined below *DD*, the price would fall. But it would fall less if the sellers were in a position to reserve their stocks, as indicated by *SS*, than if the amount offered for sale were not reserved.

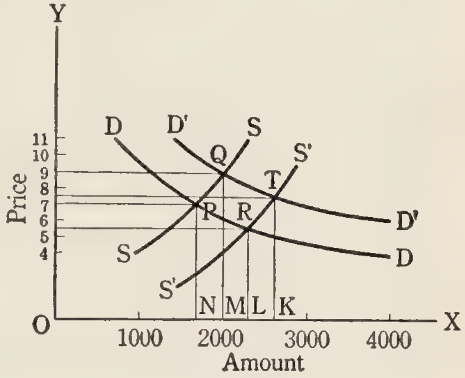

FIG. 12

Effects of Changes in Supply; Demand Constant. An increase in supply is represented by *S'S'*. If demand remains at *DD* when this change in supply takes place, price will decline to *LR*, or 5½ cents. The amount sold will increase by the amount *NL*. Again it is self-evident that if supply should decrease, that is, if *SS* should move to a position above and to the left, the price would rise above *NP* and fewer than *ON* units would be sold.

Effects of Changes in Both Demand and Supply. If demand and supply both change, as represented by the movement of *DD* and *SS* to positions *D'D'* and *S'S'*, the new price is *KT*, or about 7½ cents, and the amount sold is *OK*, or 2600 units. When both demand and supply are changing, the effect upon price depends on the extent to which they change with reference to each other. If demand increases more than supply, price will rise; and if supply increases more than demand, price will fall. These are the results that tend to follow in well-organized, highly competitive markets. When competition is very imperfect and when the buyers and sellers have inadequate information about market conditions, great changefulness of demand and supply leads to chaotic movements of price. It is possible for an increase in supply that is much greater than the increase in demand to be accompanied in portions of the market by a rise instead of a fall in price, or at least a decline in price that does not result in a cleared market. When this happens, the period of high price (high relative to actual conditions in the market) is followed by a compensating decline in price below the equilibrium point. But such conditions are self-correcting. The sellers finally discover that they cannot dispose of their stocks when price is too high and that they cannot supply the market when it is too low. Hence price moves more or less rapidly in the direction of the equilibrium point.

Marginal Buyers and Sellers. In the price-making process certain buyers and certain sellers are likely to be in critical positions with respect to their influence upon total supply and total demand. In the following table four purchasers are ready to bid different prices for one unit of a commodity, and four sellers having one unit each are ready to offer the good at different prices.

A will buy 1 unit at $5.00	E will sell 1 unit at $2.00
B will buy 1 unit at $4.00	F will sell 1 unit at $2.50
C will buy 1 unit at $3.00	G will sell 1 unit at $3.00
D will buy 1 unit at $2.00	H will sell 1 unit at $4.00

Only three units of the good will be sold at a price of $3, for at that figure the total amount the buyers stand ready to take is exactly equal to the total amount the sellers stand ready to offer. The lowest bid of $2, by D, and the highest offer of $4, by H, cannot be satisfied. It will require a price of $4 to bring H into the market, and at that figure four units will be offered and only two bought; E, F, and G will cut prices, because they stand ready to sell at less than $4;

and when price reaches $3, buyer C will come in, but seller H will have dropped out. Likewise it can be shown that if the price is started at $2, competition among the buyers will quickly eliminate buyer D. In this illustration C is the marginal buyer; that is, he is the buyer who is most sensitive to a rise in price above the equilibrium point of $3. He is at the margin, or limit, of the demand which can be satisfied in this market, and any rise in price will cause him to drop out. In like manner G is the marginal seller, because he is at the margin of the supply, and any decline in price would cause him to withdraw.

It must not be supposed that either C or G determines the market price by setting his bid or offer at $3. If either A or E withdrew from the market, a new price would be established just as it would if the marginal buyer and the marginal seller withdrew. But neither A nor E is likely to withdraw, because the former is getting for $3 a good for which he was prepared to pay $5, and the latter is selling for $3 when he would have taken $2 rather than not sell at all. The portion of the supply controlled by G is, therefore, that part of the total supply which is most likely to be cut off if the price rises.

When the total demand for a good is made up of many bids, and when the total supply is composed of many offers, it is the marginal increments of demand and supply upon which attention must be centered in attempting to forecast price changes. In Fig. 12 the marginal increment of supply (the supply being represented by *SS*) is that amount which is put upon the market with a reservation price of exactly 7 cents. This is the portion of the total supply which would be withdrawn if the price decreased below 7 cents. The marginal increment of demand is likewise the portion of the total amount demanded which buyers would withdraw if the price increased slightly above 7 cents.

The marginal supply price of any commodity is the price at which the marginal increments of supply are put on the market, and at the same time the price which is necessary to call forth these final increments. The marginal demand price is the price at which buyers stand ready to buy the final increments actually purchased. It is the price which the buyers think it just worth paying for the final increments purchased. In a perfect market the marginal supply price and the marginal demand price exactly coincide. If the marginal increment of demand proceeds from a single buyer, then the buyer is the marginal buyer. But many buyers may participate in the marginal increment of demand, and all of them may reduce their purchases slightly if the price rises. Similarly the marginal increment of supply may be furnished by a single seller or by several sellers.

Equilibrium Prices and Actual Prices. In the preceding sections the theory of market price under complete and free competition has been explained. In this explanation certain assumptions have been made concerning the conditions under which the selling and buying take place. All these assumptions are realized only if the most highly perfected marketing organization prevails; but if the organization is not perfect, some of these assumptions will fail and consequently actual price will vary from the equilibrium price. These assumptions must now be examined. We have assumed that competition is entirely free, that each seller and each buyer follows his own judgment and his own desires in determining what he will ask or bid for a commodity. This does not mean that he is entirely isolated from and ignorant of the actions of other buyers and sellers, but that there is no combination or collusion among buyers or sellers for the purpose of coercing those on the opposite side of the market. The extent to which such competition is realized in actual markets depends upon the type of market, — whether a consumers' or producers' market, — the kind of good dealt in, the number and size of the concerns buying and selling, legal restrictions upon production, such as patents and copyrights, and upon the extent to which the production of the good involves or does not involve the utilization of limited natural resources.

In consumers' markets there is almost complete competition among the buyers, for only in rare instances is it possible for consumers to combine against the dealers. The buyers are too numerous to combine readily, and the time and effort necessary to make such a combination effective are too great. On the other hand, there is always a likelihood that the dealers selling the same good in any locality will be relatively few in number, that they will always be in more or less close contact with each other, and that they will be able to effect an understanding in regard to prices. Such understandings range all the way from monopolistic arrangements that enable the participants to control the supply very effectively to rather ineffectual understandings that it is contrary to good business policy to "spoil the market" or to "do business at less than a fair customary profit." The effect of such combinations is to raise the equilibrium price which will be established above that which would prevail under the assumption of free competition. A practical example of the effect of loose combinations upon prices is often found among the merchants of a small town. As many as six or seven stores may serve a market demand that could be more cheaply supplied by one or two, and yet no dealer will try to drive out his competitors by cutting prices.[1] Such combinations raise the supply curve above its competitive level. At the same time they cause it to be more nearly horizontal; that is, they tend to stabilize the supply schedule and make it uniform at a "fair" figure, regardless of the amount sold. When the supply schedule is so stabilized, a change in demand will not produce a change in price; for if the supply curve is a straight line, it is evident that variations in demand do not affect price.

In markets for producers' goods combinations among sellers may frequently be met by counter-combinations or understandings among buyers. Hence the demand curve may be artificially affected as well as the supply curve. The term "artificial" must not be construed too literally. It is quite as

---

[1] There are, of course, many other causes of inefficiency in merchandising.

"natural" for either buyers or sellers to combine, if they find it feasible and advantageous to do so, as it is for them to compete. The fundamental human trait which is constantly at work in economic dealing is the desire to obtain the largest possible profit for what one has to sell and to give up the least possible of one's goods and services for things one wishes to buy. If it appears that this end can be most advantageously accomplished by combinations with other sellers or other buyers, men will usually or typically try to bring about such combinations. But the economist contends that such combinations are usually difficult to perfect and to keep in operation. When prices are so high that all sellers receive very large returns from their transactions, it is obviously advantageous for any member of a combination to expand his sales. If shoes which cost five dollars to put on the market are sold for six dollars, every seller is under the strongest sort of incentive to expand his sales. But in order to bring this about he must get business away from his fellow dealers, and to do this he must usually offer the buyers some economic inducement. There are a number of ways of accomplishing this. He may either reduce the price of the shoes outright or provide certain services, such as liberal credits, better salespeople, more variety, or a more liberal policy in respect to unsatisfactory goods. An actual price which is above the competitive equilibrium price is therefore almost certain to be "nibbled at" by ambitious and aggressive sellers who see greater advantage in expansion of sales than they do in the maintenance of high prices with less volume of trade.

The second important assumption implicit in the theory of equilibrium price is that of complete knowledge of market conditions. The term "market conditions" is necessarily general and inclusive. Some specific examples may, however, be discussed briefly. If some of the sellers in a market are not aware of the prices others are asking, we may find that instead of one market price there are several. Such a condition cannot prevail in respect to important commodities, however, unless the buyers are likewise unaware of the prices being asked by different dealers. Otherwise they will all desert the high-priced sellers and purchase from those whose offers are lower. But many "convenience goods" are sold at varying prices in urban retail stores within a radius of a few blocks because both buyers and sellers are lacking in complete knowledge of the prices paid in the various establishments. Price differences also arise when the buyers are unable to identify and compare the articles sold by different dealers. Thus a dealer in shoes who sells unbranded goods and who has built up a large good-will among the buyers in his market may sell somewhat above the figure charged by other dealers. Consumers cannot afford the time necessary to make accurate comparisons between the wares offered by different merchants, and in many cases they have not the necessary technical knowledge to make such comparisons. In the case of branded goods or goods whose quality is discernible at a glance by buyers of ordinary experience, prices are more likely to be uniform. Nor are the variations in price which have just been described likely to be large; but if they are, they are unlikely to be permanent. After all, the one most significant fact in retail merchandising is price, and sooner or later the lower prices asked by competitors will have their effect upon the business of the firm that attempts constantly to take advantage of the consumer's lack of knowledge.

How closely, then, do actual prices correspond to the competitive prices explained by the theory set forth in the preceding sections? Is the theory merely a quite logical conclusion from unreal and unwarranted hypotheses? Many critics of the theory fail to observe that in retail stores goods are practically never sold simply as physical objects. Some service is performed by the seller in connection with the objects vended. This is only to repeat the familiar statement that production is the creation of utilities. A pair of shoes sold by an efficient salesman in a reliable shop that permits the consumer to pay at the end of the month, that makes good without quibbling all obvious defects in merchandise, and that delivers such wares at the consumer's residence — such a pair of shoes cannot be compared in price with another identical pair sold in a shop where some of these productive services are not available. The consumer buys services as well as material goods.

Yet when due allowance has been made for these merely apparent discrepancies, there remain many instances in which the actual prices are different from theoretical competitive prices. But they are not widely different. The same quality of shoes does not sell for prices differing by 20 per cent in different shops in one locality when those shops offer substantially the same services. Nor are the differences easily maintained. Sellers may, by constantly inventing new methods to draw trade to their shops and by contriving new devices for concealing quality, continue to keep their prices above the most common price. But just as is the case with combinations, these devices soon become outworn and always tend to lose their effectiveness. All that the economist contends, or ought to contend, is that the equilibrium price is the point toward which actual prices always tend. In many markets and for many sorts of good the approximation is close; in others there is constant deviation from this theoretical price.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Criticize the following statement: The economist's theory of market price applies to the highly organized markets, such as the produce exchanges. It does not assist in understanding the price made in retail markets.

2. A branded good which is also widely advertised sells for exactly the same price in all stores throughout the country. Is this a competitive price? Is it an equilibrium price? What would probably happen if the volume of sales was too small to take the output of the manufacturer?

3. Criticize the following statements. Which are correct and which incorrect?

a. If price rises, demand falls off.
b. Demand depends on price.

c. Price is partly determined by competition among sellers.

d. Price is partly determined by competition among buyers.

e. When there are many buyers and only one seller, price is not determined by demand and supply.

✓ **4.** The extent of the change in price that will follow from a change in supply depends upon the elasticity of demand. Why?

5. What conditions or events are likely to bring about a change in the demand for wheat? rayon? automobiles?

6. Are reservation prices attached to the annual wheat crop by farmers? to men's straw hats by retailers? to old paintings by dealers?

7. In villages we often find that there is only one dealer in each line of retailing. Yet the prices which he charges are substantially the same as those charged by other sellers in neighboring towns. Why does the dealer not charge a price which is considerably above the competitive price? Is the supply which he furnishes the only supply in the market?

8. Do you think that the annual wheat crop of the United States sells at a price that yields the maximum gross receipts to the farmers? Why, or why not? If not, farmers would gain by destroying a portion. Why does such destruction practically never occur?

CHAPTER IX

CONSUMERS' DEMAND AND UTILITY

The Analysis of Individual and Collective Demand. In the preceding chapter the demand schedules of the individual and of the entire market were employed in the explanation of market price. The statement that the more of a good there is to be sold the lower will be the price that people will pay for it was relied upon to explain the determination of market prices. But we did not fully explain *why* individual and collective demands display this tendency. Moreover, no explanation was offered for the fact that the prices which consumers stand ready to pay for some goods are high while those that they will pay for others are low. The explanation of market price in terms of demand and supply remains valid even though we should be unable to discover and describe the forces that determine the demand of consumers for finished goods. Observation has amply demonstrated the important characteristics of the collective-demand curve, such as its downward and outward slope, its tendency to vary, and the varying elasticity of demand for different goods.

In all discussion of the forces back of consumers' demand the economist must proceed with caution. At no point is a final answer to the problems of demand which we have enumerated in the preceding paragraph in sight. For reasons that are obvious the conduct of buyers in the consumers' market is as yet only partly understood. Some of the statements we shall make are therefore hypothetical. But these hypotheses are not wholly unverified, and they are the most satisfactory explanations of the facts that have thus far been advanced. At present the verification of these hypotheses depends upon three sorts of data: (1) We can rely, to some extent at least, upon the statements of those who are engaged in selling goods to consumers. (2) We can bring to bear upon the problem the findings of other sciences that deal with human behavior. (3) Part of our proof consists of general observations that any person can verify for himself by noting the actions of consumers. In testing the conclusions of economics about consumers' demand three questions should therefore be constantly kept in mind: (1) Do these conclusions agree with the statements of those who deal with consumers in the market? (2) Are they in accord with the accepted findings of psychology and sociology? (3) Are they in accord with everyday experience?

It has been said that in studying the demand for consumers' goods we are dealing with human behavior and with human motives. It might seem, therefore, that the subject is primarily one with which psychology is more competent to deal and one which really lies within that field and not within the field of economics. But this is not necessarily the case, for economics and other sciences besides psychology are concerned with human behavior. The historian, the lawyer, and the criminologist cannot refuse to consider it even if they would. The investigation of human activities and their causes cannot, therefore, be supposed to be the exclusive domain of psychology. The economist must, of course, accept the findings of psychology as to man's fundamental traits and his fundamental modes of behavior. But not all these findings are relevant to a study of economics. Economics studies only a portion of man's activities — only a limited division of the entire field of human behavior and human motives. It professes to study only those activities having to do directly and indirectly with buying and selling goods and services. As far as the subject of the present chapter is concerned, we shall deal primarily with the exchange of goods for money. Yet it should be noted at the outset that the student of economics is aware that it is impossible to separate economic conduct from other activities of the individual. No individual is entirely an economic man any more than he is a political man. But because we are concerned with buying and selling — with the problems of the market place — we must arbitrarily separate, as far as we can, man's economic conduct from his political, religious, and other activities. Such a separation is an abstraction, but it is a type of abstraction that is common in other sciences. Clearly, then, it is no criticism of our method to point out that under the stress of religious motives men sometimes give away their goods, work without wages, and devote their lives to the service of other people. We are not engaged in a study of men acting under the dominance of religious or philanthropic motives. But since men *are* moved by considerations other than economic, the findings of economics must, when practical activities are contemplated, be checked by the conclusions of those who are engaged in the study of other types of human activity. Economics only undertakes to explain economic behavior.

In studying the behavior of man in earning his living the economist is justified in making certain assumptions which are typically true for the business man, the wage-earner, the investor, and the consumer in their *economic* activities. We assume that the consumer always seeks to secure the maximum gain for the least cost, whether that cost is in terms of goods or in terms of the sacrifices of labor-effort and saving. We assume that in his economic activities the consumer, the seller, the wage-earner, or the investor will act according to the knowledge at his command; that he will make intelligent decisions and choices in respect to the goods he buys, the business he conducts, the employments he selects, the loans he makes. When we say that men act intelligently, we mean that they take into consideration the known facts relating to the market, that they compare goods and prices, and that they follow the course of action that leads to the maximum gain as they see it. We do not mean, as economists have sometimes been supposed to believe, that men are mechanical calculating machines constantly engaged in weighing pleasures and pains. We assume that since men are in business they will try to make the largest possible gain for a given outlay of money and effort — not the immediate maximum, but the highest consistent with the maintenance of that gain for a considerable period of time. We know, of course, that none of these assumptions are true of all people at all times. This economics does not assert. It does insist, however, that the assumptions are true of the great majority of men in their economic activities.

Utility. All things that a consumer consciously and voluntarily chooses to buy must have for him some personal significance; that is, at the time of choosing and buying he believes that they will satisfy some personal want. These things are the objects of desire, and they stimulate or produce in the individual the tendency to acquire them. Stated briefly, *utility is the capacity of a good to satisfy human desires.*

Much needless confusion has arisen in economics over the implications of the term "utility"; and if the student is to avoid this confusion, it is necessary that he should be made aware of its causes at the outset. In the first place, the word "utility" carries no moral significance. If at the moment of purchase a good gratifies a desire, it has utility, no matter whether it confers a lasting benefit upon its buyer or inflicts a permanent injury. Thus the narcotic has utility for one who has become habituated to its use, just as it has for the sick person for the relief of whose suffering it is prescribed by the physician.

In the second place, the want which is satisfied by the good 2 possessing utility is not necessarily a sensuous one. We desire many things that do not give us pleasure in the narrow meaning of the term. We want books, not only because they amuse or delight but because we believe that they will instruct, because the habit of reading has been developed in the schoolroom, or even because a friend has suggested to us that we ought to read them. We frequently wear a certain kind of clothing because it is the fashion or the custom or because we wish to avoid the criticism of others who expect us to wear it. Nevertheless these goods are the means by which we satisfy a want or accomplish certain desired ends.

In the third place, we must keep in mind the fact that 3 *utility* pertains to individual reaction. It is the power of a good to satisfy the wants of an individual, and not the abstract and general significance of a commodity for the comfort or well-being of the community at large. This statement does not deny that individuals are controlled by the action and opinions of larger groups. They buy goods because their friends buy them and because they are subjected to the same advertising, the same climatic environment, the same general social influences as their friends. With all the similarity of reaction to these stimuli there are, however, individual variations that are of importance to economics. Hence we must refrain from using the term "social utility" to mean the want-satisfying power of goods for individuals taken collectively as members of society.

The utility of a good is indicated and measured by the individual reaction to that good. The particular qualities which cause the good to produce in the mind of the consumer the desire to acquire it are partly controlled by his biological make-up, partly by environmental conditions, and partly by acquired modes of reaction. Bread has utility for a large portion of the race because it satisfies a biologic need. The utility of fur coats in winter and of light clothing in summer depends upon changing conditions of the weather. The utilities of many other goods, including things which serve no obvious physiological need, are controlled by the social environment rather than physiological need. Our notions of what is beautiful in the decoration of our houses, of what is desirable as recreation, or of what is in good style in our clothing are controlled very largely by the opinions, customs, and fads of the community in which we live. Utility is not innate in the good itself, nor is it determined by a changeless attitude of the consumer toward his natural environment. Men are not furnished at birth with ready-made utility curves. Education, the appeal of advertising, the dictates of fashion — all operate to modify the utility that goods have for consumers.

Finally, the term "utility" can be *strictly* applied only to goods intended for personal consumption by the buyer or by someone for whom he is acting. Thus cloth has utility because it ministers to a human want, but the machine on which it was woven has none. Raw materials, the larger proportion of land, and other producers' goods have no specific utility. It is sometimes said that such goods have *derived utility* because they aid in producing finished goods that do have the power to satisfy wants. But this terminology may lead to confusion. It can hardly be asserted that a machine used in fashioning one of the parts of an automobile has utility to the users of the motor cars it aids in turning out. The users are generally unaware that such a machine exists; and even if they do know about it, they do not desire it. To them it is a matter of indifference what machines or what materials are used to make a car, as long as it satisfies their desires. It is also inaccurate to say that the machine has utility for the owner of the manufacturing plant. He may "want" it because it performs certain work which in turn enables him to make a profit on his investment. But the basis of his demand for the machine is a cost-profit calculation and not a calculation involving his personal satisfactions. Clearness demands that we should not apply the same term, "utility," to two different phenomena. It is better simply to say that the machine is of economic significance to the manufacturer. He wants it because it aids in the production of utilities.

Utility and Scarcity. It is a matter of common observation that goods that can be had for the mere trouble of taking them have no value and no price. But this condition cannot exist unless these articles are so abundant that all may consume them at will. Such objects are called free goods. On the other hand, all goods that do not exist in large enough quantities to satisfy all who desire them are called economic goods. Both types have utility, but only the latter has value. The relation of scarcity to value seems at first glance a simple matter. The occurrences of everyday life lead us to conclude that unless a consumers' good has both attributes, scarcity and utility, it cannot have value. For many years economists were content to stop with this conclusion. But this statement did not explain why some goods are more valuable per unit than are other goods, why the consumer is willing to pay a high price for a yard of silk and only a low price for a yard of cotton, or why high prices for both goods prevail at times and low prices at other times. It is also true that a brief experience in selling any commodity will soon convince one that as a rule increasing amounts can be marketed only if the price is lowered. But the market does not offer any explanation of this inverse relation between price per unit and quantity marketed.

The Principle of Diminishing Utility. The inverse relationship between price and quantity brought to market depends upon the principle of diminishing utility. Since utility is an individual matter, we must begin our explanation with the reactions of a single individual who acquires varying amounts of a consumers' good. Let us assume that the good is a necessity of life, such as bread, for which there are substitutes but none entirely satisfactory to those who have become accustomed to its use. If the quantity annually produced were much more limited, many persons would be compelled to curtail their use of it. And if the quantity available for consumption were gradually diminished, the loss would be felt more and more by each individual. With present abundance no one attaches much importance either to the acquisition or the loss of such a small quantity as an ounce. But if everyone were reduced to a siege ration of one or two loaves a week, the loss of any given ounce would mean a sensible deprivation. Let us suppose that each individual has only four ounces a day for consumption. If one ounce more is added to his allowance, he will feel a certain gratification. He will find that his desire for bread is more nearly satisfied. If a second ounce is given him, his satisfaction will be still more nearly complete; but the second ounce will not seem so important to him as the first that was added to the existence allowance of four ounces. The utility of a unit of bread is thus seen to depend upon how much is available for consumption during a given period of time. The want which any given unit satisfies is more or less intense, depending on how many other units have already been consumed or are available for consumption. And the relation between utility per unit and quantity is seen to be an inverse one.

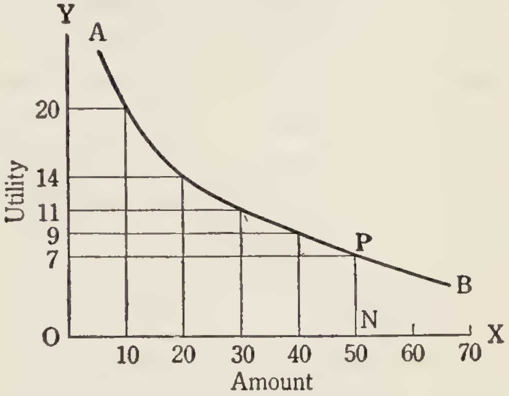

FIG. 13

Without asserting at this point that the principle is generally applicable to all consumers' goods we may state the law of diminishing utility in respect to consumers' goods as follows: *As the* ~~*quantity of a good acquired for*~~ rate of *consumption increases, the utility of each unit acquired is less than that of the preceding unit.* Strictly speaking, this is not a law of *diminishing* utility but of *the inverse relation of utility per unit and amount available for consumption.* This relationship may be illustrated by the curve *AB* in Fig. 13. Along the axis *OX* are measured the amounts available for consumption; and along *OY*, the utility per unit. If only ten units are available, the utility will be high; but if more units are added, the utility of each will become less and less. In the figure definite arithmetical values are given to the ordinates extending from *OX* to the curve *AB*. This involves inaccuracy, *because there is no method of measuring utility except in terms of something which the individual* *will sacrifice to obtain a given commodity.* The *absolute amount* of utility might perhaps be measured accurately if we were able to gauge the varying mental reaction of the consumer by some purely objective standard, as we can his purely muscular reactions to an electric current of varying intensity. Yet the measurement of his response to the good in question in terms of money is certainly not grossly inaccurate. The amount of money a consumer spends for any one good (not an entire class of goods, such as food) is as a rule so small a part of his entire income that the expenditure does not make each dollar more significant than it was before. As Alfred Marshall has put it, we may assume that the utility of money remains practically constant. If now the consumer is willing to sacrifice less and less money for items of bread as he is supplied with more and more of it, the inference that this is true because it is the utility of bread that is declining cannot be much in error.

The proof of the law of the inverse relation of utility and
amounts available for consumption rests primarily, then, upon
1 the observed fact that as the individual extends his consump-
tion of a commodity his willingness to pay money for each unit
declines. In the second place, statistical studies of the demand
2 for goods which are used almost entirely for human consump-
tion, such as potatoes, show a pronounced decline in the
amounts of money people are willing to pay for a unit when
the total amount to be sold is increased. Finally, it is a matter
3 of common observation that the more of a good a person has
the less does he treasure any particular item of it.

It is sometimes asserted that the principle of diminishing utility applies only to "dinner-table examples." But this criticism is without foundation. The use of all consumers' goods, whether articles of food, or clothing, or housing, or amusement, exhibits this same phenomenon of variation in utility per unit as the quantity varies. There are, however, certain limiting conditions attaching to the statement of the law. A definite period of time must be set within which the consumption of the good takes place or within which its use is contemplated. If the individual consumes six ounces of bread today, an additional six ounces acquired for use tomorrow will have the same utility. This the law does not deny. It only asserts that if he adds another ounce to his daily ration this additional quantity will have less utility than the preceding amounts. In the second place, the individual must retain the same attitude toward the good in question. If his tastes change, additional units of the good may come to have either a larger or a smaller utility. Finally, there must be no change in the environment which will affect his desire to consume. With the change of seasons from autumn to winter the outdoor worker finds that the utility to him of a given additional quantity of food increases.

Marginal Utility. As the amount of a good acquired by the consumer increases, the utility of each unit declines. If the stock (Fig. 13) is limited to fifty units, the utility of the last unit will be represented by the ordinate *NP*. Since the fiftieth, or last, unit has a given utility measured by this ordinate, that degree of utility is called the final degree of utility or the marginal utility of the entire stock of fifty units. Since the units are by assumption interchangeable, no unit of the stock can have either a greater or less utility than this last, or marginal, unit. To prove this statement let us assume that one of the units other than the last one has been lost or destroyed. How great will be the deprivation suffered by the consumer? If the units are alike and interchangeable, he can put the last unit in the place of the one lost. Hence the loss is measured by the utility of the last, or marginal, unit. The effective importance of the entire stock is measured from the utility or importance of the last unit; what is called the *effective utility* of the entire stock is found by multiplying its marginal utility by the number of units consumed or acquired. The effective utility of any stock is distinguished from its *total utility.* If a monopolist could dole out to the buyer the stock of fifty units one by one and at the same time keep him in ignorance of the existence of any further supply, then each unit would be appraised by the buyer at the utility measured by the corresponding ordinate extending from OX to AB. The first few units would have an extraordinarily high utility, the tenth a lower utility, the thirtieth a still lower utility, and so on. The sum of these utilities would be measured by the area between the axis OY and the ordinate NP and between the curve AB and the axis OX. Total utility is evidently a less definite and measurable quantity than effective utility. No one knows how great is the utility to him of a very small quantity of many of the commoner necessities of life, and to construct a curve showing the total utility of such articles is therefore impossible. With less necessary articles of consumption the case is somewhat different. Most of us are quite aware that we would not buy ice cream at a dollar a pint, because the utility of even one pint is less than the utility of many other things a dollar would purchase.

Thus far we have considered the application of the principle of diminishing utility to goods of which the consumer buys a large number of interchangeable units. But many things are not of this sort. We do not buy unit after unit of pianos, or houses, or automobiles. Do the phenomena of diminishing utility and marginal utility apply to such goods? Commodities such as bread satisfy very simple desires. Other goods, such as automobiles, cater to complex desires. The latter good may be desired because it provides comfortable transportation, because it furnishes the owner with a kind of recreation, because it ministers to the desire for distinction, and for many additional reasons. Now the utility of a good depends, of course, upon the services it renders, upon the desirable events that, in the opinion of the user, flow from that good. This is true of automobiles as well as of bread, but in the case of bread the service is brief, the commodity quickly consumed, and the want simple. In the case of the automobile the services are numerous and continued over a considerable period of time. Hence the buyer of the latter good usually considers the desirability of each of the several services and also the repetition of those services. The principle of diminishing utility applies to the repetition of each of the services of the automobile. For example, a consumer might well ascribe a high utility to a small amount of the sort of service rendered by an automobile, but more than a small amount might have for him a very low utility. If he greatly desired a small amount of automobile service, he would probably hire a taxicab rather than buy a car outright; but the phenomenon of diminishing utility would still apply to his consumption of automobile service. Because the service derived from bread is quickly rendered, the identification of the service with the physical object is close and we do not observe that the two are different. But in the case of the motor car the difference is easily seen. In both, however, the utility is derived from the use of the good, and a given amount of service from either would have for a consumer a given marginal utility.

Utility and the Individual-Demand Schedule. It is the purpose of this section to explain how the consumer arrives at the demand prices he sets upon different amounts of the various goods he buys. A great majority of consumers do not receive incomes sufficient to gratify all desires. Many objects in a consumer's environment attract his attention and produce in him a reaction which disposes him to use or possess them, but he has not the wherewithal to acquire them all. Since the consumer cannot gratify his every whim, *he must make choices among goods,* acquiring some of them and forgoing others. For purposes of illustration let it be assumed (1) that there are only three goods among which an individual can at a given time make choices and (2) that he has not the means to acquire all he wants of all three. On what basis will he make his choices? Which of the three will he acquire? Or if he acquires some of each, what will govern the amount of each that he acquires? The answer to these questions is that if the consumer is rational and carefully considers the utility to be obtained, he will make such a selection among the goods as will give him the greatest satisfaction. He will be attracted first to the good which has for him the greatest utility. If his income is very limited, he may spend it all in acquiring this good and nothing else.

In order to explain the choice of different amounts of several goods we may suppose that the consumer has the power to buy 100 units of bread or 80 units of meat or 60 units of pastry. But if he acquires much bread, he will find that its utility declines as the amount is increased. The hundredth unit may have very little significance, and consequently the total stock will have a small effective utility. Under the circumstances he will turn to meat and even to pastry before he acquires 100 units of bread. As he increases his stock of each, the utility of an additional unit of each will decline; and when he has entirely expended his means of acquisition, each good will have for him a given marginal utility. If he is rational and if he carefully considers the marginal utility of each good, it will be found that when he has expended all his income he will have so distributed his consumption that, *relative to the prices he has to pay for the various goods*, his marginal utility for all will be equal. The marginal utilities will not be *absolutely* equal. If the prices of bread, meat, and pastry are 10, 40, and 60 cents a unit respectively, then the marginal utilities of the three goods will stand in the relation of 1, 4, and 6. The consumer would not continue to buy meat to the point where its utility, relative to bread, was as 2 to 1, because he would then be spending money upon meat that would procure for him a greater satisfaction if expended for bread. He would not continue to consume pastry to the point where its marginal utility was only four times that of bread, because at that point he would be compelled to pay for a unit of pastry 60 cents, which would buy an amount of bread having more utility per unit. The prices which the consumer tends to pay for goods are thus seen to be proportional to the marginal utilities these goods have for him.

If we bring the illustration nearer to real life by introducing many commodities, the explanation will be in no wise altered. Instead of comparing the marginal utilities of three goods, the consumer would seem to be confronted with the problem of comparing the utilities of a large number of goods. But in deciding whether to buy a given quantity of a certain commodity he does not need to make the multitude of comparisons suggested in the preceding sentence, provided (1) money is in common use and (2) he has established a scale of expenditure in which he can place with more or less accuracy the utility of any good with which he is familiar. He knows that if he buys too lavishly of one good, he will have to go without a number of other things which would have for him a greater power to satisfy wants. He tries, therefore, to distribute his income so as to obtain the maximum return in utility. In a figurative sense money comes to have a definite utility per unit, because it comes to stand for a composite of utilities derivable from a large number of goods. The consumer therefore tends to continue his purchases of a particular good until the marginal utility he obtains from it is equal to the utility represented by the amount of other goods he could purchase with the money the last unit of the particular good costs.

In the case of durable objects, the consumer compares the succession of uses, pleasures, or benefits he will receive from them during a period of time with other satisfactions their cost in money would buy. Here again he will tend to lay out his income in such a way that the money spent for such goods will yield him about the same marginal utility as could have been purchased if the money had been used to buy other commodities. The preceding statement is subject, however, to the qualification that the utilities to be derived from the use of a permanent good must ordinarily be greater than the utilities of an impermanent good having the same price. This is true because the satisfactions derived from the durable good are in part deferred to the future, and it is a common rule of life that future satisfactions are undervalued as compared with those of the present.

In the preceding discussion the reservation "if he [the consumer] carefully considers the marginal utility of each good" has been employed. The meaning and significance of this reservation must now be explained. If the buyer should purchase at random, paying any price that was asked, or if he should never stop to consider what any good cost him in terms of other goods given up, or if he should act blindly and on mere impulse, then for economics his conduct would be irrational. Under these conditions his choices of goods would have nothing to do with utility. In place of making his choices on the basis of comparative utilities, he would make them on no basis whatever, and utility would be irrelevant. But under the conditions of buying and selling as we find them the consumer does make choices and he does weigh the desirability of one good, or a given number of units of one good, against varying quantities of other goods. The proof of this statement is found everywhere in the market. The art of consumption is far from being exact, but the practice of it is not as chaotic as is sometimes represented. Automobiles that do not "stand up" cannot be sold; stoves that do not heat as well as other makes are discarded; clothing that does not wear well or that proves unsatisfactory in other respects remains on the shelves of the dealers.

In some instances the consumer may be misled. The expert can show that he has been deceived and that he really chooses the inferior good. But this is beside the point. Regardless of the wisdom of the consumer's selection, he has made a rational choice. It has also been objected that the consumer cannot compare the utility of literally thousands of articles offered for sale in the shops of a large city. This is partly true. It is especially true in the case of goods that the consumer buys only infrequently and concerning which he has therefore only a limited experience, or in the case of goods that are new to him. In these instances the shrewd buyer will usually do one of two things: either he will patronize dealers who have established reputations for honesty and fairness, or he will try to obtain the opinion of one who has had more experience — the expert, perhaps. Thus one buys a new accessory for his automobile on the opinion of other owners who have used it, on the advice of a mechanic, or on the recommendation of a reputable dealer. In any case it is clear that in this instance the user of the article is partly substituting the choice of another for his own judgment; but only partly so, for no matter how highly recommended the good may be, the prospective buyer may still conclude that he cannot afford to buy it. That is, he decides that on the representation of other people the utility of the good is not great enough to warrant the expenditure of the money necessary to obtain it. The major part of one's income is not expended for new goods or for goods with which he has had no experience; in most cases the buyer has used the identical good or has used one sufficiently like it to provide a basis for comparison.

Goods commonly consumed are, however, so numerous that the buyer cannot stop every time he makes a purchase to compare its marginal utility with the marginal utility of money — with the utility of the other things its price represents. Fortunately he does not need to do this. Any experienced consumer soon establishes a scale of expenditure in which he has placed roughly the relative importance of quantities of various sorts of goods. The places of different goods in this scale depend on a great number of considerations: income, education, the acts of associates, climate, advertisements, and many other conditions. The net result of this ranking is that the consumer comes to know how much a dollar is worth to him in terms of given quantities of a great variety of goods. He knows without elaborate calculation that if he buys a pair of shoes for double the price he is accustomed to pay, or if he buys two pairs of shoes instead of one pair, he will be compelled to go without certain other goods whose sacrifice will cause him annoyance or disturb unpleasantly the routine of his life. If the purchase is an important one, he will stop to consider these sacrifices. He will weigh as well as he can the various returns in satisfaction he might obtain from each good. But numerous small purchases he will continue to make at approximately the same price and in the same amounts without giving either much thought.

In the case of goods that cannot be bought in small units or whose services cannot be obtained in small units the consumer may find it impossible to arrange his budget of expenditures so that the marginal utilities of all the goods he purchases will be proportional to the prices he pays. For example, an individual might wish for radio equipment for the purpose of listening to programs of a certain sort, — say, reports of football games, — but he might not attach a high utility to any additional use that he could make of a radio. He wants the equipment for use on only six or seven afternoons during the year. For this service he might be willing to pay $10 a year. But he finds that the annual cost of such an outfit as he would require would be $20 a year. The additional services he could get from it are not worth $10 annually. What should he do? If he buys the radio, he will have to forgo some other service or commodity which has a cost of $10 and a higher utility than the additional service of the radio. If he does not buy the radio, he will probably spend $10 on some other service or commodity which will not give him as much satisfaction as he could obtain from listening to the reports of football games. He may be able to strike a balance between the excess of satisfaction derived from the reports of games and the loss of alternative enjoyments resulting from paying $10 for the additional radio services and giving up other more important satisfactions. But if the reporting of games is only just worth $10 to him, he will not be able to do this. Then he will have to make up his mind to losing potential satisfaction either way he decides. He will be unable to derive the same satisfaction from marginal expenditures on all the commodities and services he buys.

The individual-demand schedule represents the prices that an individual stands ready to pay for different amounts of a given good. From the preceding explanation it will be seen that these prices will be paid by the individual only if (1) his money income remains the same, (2) if his demand schedules for other goods remain the same, and (3) if his desire for the given good remains the same. If his income shrinks, he will need to revise his scale of expenditures for all goods. Hence he will not be willing to take as much of any good at a given price as he would have taken before his income declined. Likewise if the marginal utility of one good is increased by a change in tastes, the consumer will desire more of it and will expend more money for it. This will necessitate a reduction of expenditures for other commodities, and therefore the consumer's demand schedules for all these other commodities will be altered to some extent.

Marginal Utility and Market Price. As we have seen, the marginal-demand prices of an individual for different goods are, in the main, proportional to the marginal utilities of the goods to him. But the question arises whether the market prices of different goods are proportional to the marginal utilities of these goods to all the buyers. The answer is that it is possible, but not necessary, that this proportion would prevail. In the first place, because of differences in income the prices which different consumers stand ready to pay for the same quantities of the same good are unlike, and the amounts that they buy of the same commodities at the same prices are unlike. The marginal utility of money is higher for the people of low incomes than it is for those who receive higher incomes. Hence the purchase of a ticket to the theater at three dollars by two people who have different incomes may be a marginal purchase for both and yet not represent the same satisfaction to both. For the one who has a small income three dollars may represent a commodity or a quantity of commodities which have a high marginal utility, but for the other it may represent goods that have very little marginal utility.

In the second place, we know that the rich buy and consume many commodities and services which the ordinary workingman's family never uses at all, and that to a less extent the latter buy and use commodities that the rich do not consume. Between these two sorts of commodities, a unit of each of which may sell for the same price, there is no necessity for equality of marginal utility any more than there would be between two commodities sold and consumed exclusively in different parts of the world. The users of the commodities are divided into noncompeting groups. In somewhat the same way extreme differences in tastes and habits of consumption may separate two groups who live in the same city and have the same incomes. There are commodities which are highly prized — have high marginal utility — by some and not consumed at all by others, because of these differences. And in this case it does not hold that the market prices of goods are proportional to their marginal utilities. The marginal buyers of different goods do not have the same scales of satisfaction in respect to different goods.

Some Practical Applications of the Marginal-Utility Theory. The utility theory, which includes the principle of diminishing utility, is the only reasonable explanation of the fact that many of the necessities of life sell at low prices and command only a small portion of the income of even the wage-earner. The world could ill afford to do without bread, and yet many wage-earners' families expend much more income for comforts and decencies than for bread. Of course the answer to this apparent riddle is that the amount of wheat that can be obtained from the grain lands of the world is so great that its marginal utility is low, whereas the sources of less necessary goods are more limited by nature, by the disutility of labor, and by saving costs, and thus the marginal utilities are high. The utility theory also explains satisfactorily the decline in price when the amount of any good offered for sale in the consumers' market is increased.

One of the most consistent tendencies of modern economic life is the constant expansion of wants. As invention and scientific discovery increase man's power to produce, we find that instead of confining production to old commodities he constantly invents new ones. Both production and consumption become more and more complex. Part of this complexity is due to the fact, stressed by Alfred Marshall, that new activities create new wants. It is also due to the principle of diminishing utility. As the old commodities are multiplied, they come to have less and less want-satisfying power; and as their utility declines, the producers, who are ever on the alert to sell goods, introduce new ways of satisfying old desires and new modes of consumption.

The effectiveness of advertising in expanding sales and in keeping up prices rests on its power to bring goods to the attention of potential consumers who would otherwise either not know that the goods existed or not have their attention constantly directed to them. In the latter case the advertisement works to produce a constant comparison between the want-satisfying power of the good to be sold, often at the advertiser's appraisal, and the want-satisfying power of the other things its price would buy. Here again the utility theory furnishes a hypothesis upon which we can explain the observed facts.

The utility theory is also helpful in such practical matters as the distribution of the burden of taxation among the different income classes of the community. The rich can be taxed more heavily, both because they have the means to pay and because the subtraction of a given proportion of their incomes through taxation does not deprive them of the same amount of utilities as would the subtraction of the same proportion from the incomes of the lower-income groups. Still another application of the theory can be made in the case of the debts owed by other nations to the United States. One hundred million dollars added annually to the income of the people of the latter country through the reduction of taxation by reason of the payment of this sum on the French debt would doubtless increase the sum total of their satisfactions. But it would also decrease proportionately more the satisfactions that would have been enjoyed by the people of France if they had not been taxed to make this payment. The typical family income of the American people is greater than the typical family income of the French. It is therefore roughly correct to say that to subtract a given sum from the incomes of the latter to increase the incomes of the former will amount to depriving them of a quantum of satisfactions that cannot be made up by adding the same sum to the incomes of the former. To be sure, the same argument might be made in respect to many debtors and creditors, and we should not on that account conclude that all debts should be canceled when the creditors are rich men and the debtors poor men.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It is sometimes said that the utility explanation of demand price is faulty because custom, whim, and the persuasive arts of salesmen determine the prices consumers are willing to pay for goods. To this we may reply: Economics does not inquire *how* the consumer got his desires, but only shows how he makes his choices, desires being given. Do you think the reply is adequate?

2. "Utility has nothing to do with demand price. If it did, people would cease patronizing the 'movies' and go to symphony concerts, refuse to buy gaudy, unserviceable clothes, and insist on quality rather than appearances in their purchases." Show that in this quotation an erroneous definition of utility is implied.

3. The utility theory is frequently criticized because, it is said, it is based upon untenable assumptions. Which of the assumptions enumerated in the latter part of the first section of this chapter do you believe are least tenable?

4. One writer on the theory of value has said that the utility of any good is regulated by the utilities of other goods. Show that this is partly correct, but that it is not correct if applied to our desires for certain types of goods. To which types does it apply? not apply?

5. The operation of the principle of marginal utility is sometimes interfered with because the objects the consumer desires are not divided into small units. Enumerate commodities that you think might be sold in larger amounts if they were salable in smaller units.

6. In this chapter it is said that the reason why people will buy increasing quantities of goods only at declining prices is the principle of diminishing utility. Assume that this explanation is incorrect. What better explanation can you suggest?

7. "The principle of diminishing utility is a mere assumption. It is inferred from the fact that prices decline as the quantity of goods offered is increased." Is this statement correct? What other evidence of the existence of the principle of diminishing utility can you give?

8. A thousand families having widely different incomes all buy milk at the same price. Is it possible that all these families are marginal buyers for the last units purchased? If so, have the quantities taken by each family the same marginal utility?

9. The elasticity of the collective demands for different goods is partly explained by the principle of diminishing utility. Why? What other economic facts affect elasticity?

CHAPTER X

COST OF PRODUCTION AND NORMAL PRICE

How Supply is Determined. The market price of any good is determined by demand and supply. Demand, as we have seen in the preceding chapter, is the result of the great variety of conditions that determine the marginal utility of a good to consumers and the purchasing power which the consumers of the good can command. Supply is proximately determined by the sellers who set reservation prices upon products and offer varying amounts for sale at the prices which they believe will yield them the maximum gain. The conditions affecting the reservation prices which sellers place on goods have been stated. To complete the explanation of price we need, however, to inquire somewhat further into the conditions affecting the relative supplies of different commodities.

It is evident that if the different amounts of consumers' goods were arbitrarily fixed by the government or some other agency, the price paid for each commodity and the relative prices of different commodities, or the exchange values of all commodities among themselves, would be determined by demand alone. On the other hand, if the demand schedules of all goods were fixed and invariable, the amount of each commodity offered for sale would determine its price, and the relative prices of different goods would also be determined by the relative amounts thrown upon the market. Now the relative amounts of different goods put up for sale are not determined arbitrarily any more than the demand for them. It has already been explained how the forces back of demand operate, and, in part, how the supplies of goods that are either completed and ready for sale to consumers or well on their way to the retail market are regulated. We have next to study the economic forces that act in a permanent and persistent manner to determine the supply prices of commodities. Another way of stating the same problem is to say that we must examine the long-run effects of economic forces, for the temporary factors affecting amounts offered for sale, and the prices at which sellers will offer goods, may be expected to cancel out if a long-time view of the market is taken.

Cost of Production and Price. The problem of normal price is, then, to discover how supply prices are regulated by permanent forces. For the purposes of simplification we shall assume that demand remains fixed; that is, that the demand schedules for all goods do not change.[1] If we should ask a manufacturer of, say, cotton cloth what he considers a normal price for the commodity he produces, he would probably reply that a normal price is one that covers his payments to other people for raw materials, labor, borrowed money, land rent, and so on, plus a fair return on his own investment and for his own work devoted to the management of the business. Of course he would also say, if he were quite frank with us, that he would like to sell at a price considerably above these costs of production in order to earn a very high return on his investment and for his own work, and that he would take advantage of any temporary shortage of supplies to charge the highest price he could get. But he would add that competition seldom permitted him to do this and that when such fortunate opportunities did come to him they were likely to be offset by the losses of bad years, so that, in the long run, he could not expect to receive a price for his product in excess of cost.[2]

If we should next ask what determines the rates at which our hypothetical manufacturer must buy the agents of production from other people, or what determines his out-of-pocket expenses, we should be answered that, as far as the single producer is concerned, they are determined by the competition of other producers who wish to employ the agents in producing cloth. And if we should ask what determines the fair return on the invested capital of the manufacturer, we should learn that each investor expects to obtain the competitive return for investments of equal risk and that no one puts his funds into a business unless assured that in all probability he will receive this return. Hence we should find that the costs of the agents of production are determined proximately by the competition of entrepreneurs within the same line of production.

---

[1] The problem of normal price under conditions of varying demand will be discussed in Chapters XIII–XV.

[2] The student's attention is again called to the fact that "cost" includes a fair return on the entrepreneur's own investment and labor.

But the land on which cotton is grown could also be used for the production of several other crops, the capital invested in machinery could be used in other businesses, and the labor used in the cotton mills might be employed in some other occupation. Evidently, then, it is not the demand of the cotton mills alone that sets the prices of these agents. The cotton industry must bid against many other industries for land, labor, loanable funds, entrepreneurship — in fact, all the agents of production. And the cost of the agents of production for making cotton cloth is thus set by the competitive bidding of all industries using the same agents of production. Hence we conclude that the cost of production of cotton cloth is determined by the value which entrepreneurs in all industries attach to the agents for producing cloth. Costs thus determined are called opportunity costs.

It may be objected at this point that we have not explained how the prices of the agents of production are ultimately determined — that we have only said that the prices of the different finished commodities tend to the level at which the agents would earn the same in all industries — and have not given any conditions that fix the values of these agents. It must be admitted that no explanation of the values of these agents has been given. But for the present discussion this explanation is not required. The ultimate determination of the values of the different groups of agents — the factors of production — is not required for the present discussion. We are here concerned to find out how the costs of production of the products of a single industry or single commodity are determined. Producers in a single industry cannot affect these costs very much. They cannot, as a rule, affect the total demand for the factors of production enough to cause a change in their prices.

Later the explanation of the values of the factors will be resumed, and it will be shown how these values are determined.[1]

Still another assumption must be made at this point. We know that while the agents of production tend to have the same costs to all entrepreneurs in the same market, the average unit costs of turning out a product are not the same in all plants. Some entrepreneurs are more skillful than others in managing production, and in consequence average unit costs vary from plant to plant. This difficulty may be put aside for the present discussion by assuming that entrepreneurs are of equal efficiency and that total unit costs are the same in all plants.

Normal Price. Since the unit price of any commodity tends toward its unit cost of production, it is sometimes said that normal price is "cost price." This statement is correct, but it throws very little light on the forces which determine the norm about which market prices fluctuate. In the following explanation of normal price we shall again resort to the use of the concrete illustration of cotton cloth. We may put the problem as follows: At what price will the supply of cloth from all the mills be such that, in the absence of temporary disturbing conditions causing market price to fluctuate about normal price, the demand for cloth will just absorb the entire annual product and no more?

The amount of cloth that can be sold at high prices is less than the amount that can be sold at low prices, because the demand curve for cloth, like all other demand curves, slopes downward. Hence, if the output of the mills were increased gradually, the price would presently fall below the cost of producing cloth. This would not be a profitable condition, and some labor and capital would sooner or later be shifted from the cotton mills to other enterprises. On the other hand, if the amount of cloth were gradually reduced, a time would come when the price was above the cost of production and the mills would make a larger money gain from the use of the agents of production than entrepreneurs in other industries could realize from the employment of these same agents. This would cause a drift of labor and capital into the cotton industry, the output of cloth would be increased, and price would fall.

---

[1] Chapters XXIV–XXVIII, on wages, rent, interest, and profits.

At some point in the upward and downward movements of the output there would, however, be a price which would neither invite any new labor and capital to enter the industry nor encourage any of these agents already in the industry to leave it. This price would be reached when the agents of production employed in the cotton manufacture earned exactly the same wages, rent, interest, and entrepreneur's profit as they did in all other industries. At this price the demand for cloth and the amount produced would be in stable equilibrium. No producer would have any economic reason for either enlarging or contracting the output of cloth. The price accompanying such a stable equilibrium of demand and production is *the normal price of any commodity*. The price of any commodity is normal if the agents of production earn as much in the production of that commodity as they do in industry at large, and no more.

Normal price is not a fixed price. Economic conditions change, and the point of equilibrium of demand and production does not remain fixed. The normal price of a pair of shoes may be five dollars in 1927; five years later, owing to changes in the opportunity costs of the agents of production, it may be either above or below that figure. Neither does the term "normal" mean "ethically right" or "socially desirable." The normal price of a quart of milk may be fifteen cents, yet it may happen that at this price the children of poor families cannot obtain enough to keep them vigorous and healthy. No injustice is practiced by the distributor who sells at this price, although there may be serious social injustice in the fact that the children are denied the requisites for healthful living. It might be entirely desirable for the state to step in and provide a cheaper supply of milk for those unable to buy at the normal price.

Prime and Supplementary Costs. In the preceding discussion it has been assumed that conditions were so stable that aside from minor fluctuations producers could predict with a high degree of certainty the outcome of their investments in materials, labor, and fixed capital. Under these conditions manufacturers would not make the error of building large mills which could not be run at capacity, nor would farmers plant tracts of land that were not required to meet existing demand for their products. We may now modify our previous assumptions by admitting the fact that demand cannot be accurately forecast and that one producer cannot know the amounts of product his competitors will put upon the market in the future.

It matters little what goods may have cost in terms of past labor, past investment of capital, and past cost of raw materials if a given quantity has been thrown upon the market. As we have seen, this quantity will have to sell for whatever the buyers are willing to pay for it. Hence it is only when past plans for production and past costs have been accurately adjusted to the demand that exists when the finished goods arrive at the consumers' market that goods sell for prices that agree with past costs of production. It often happens that entrepreneurs invest too much capital in the buildings and machinery of a certain industry. Such capital goods cannot be quickly worn out and they cannot be used for the production of any sort of commodity except the one for which they were intended. If their earnings fall below the gains that could have been obtained if the investment had been made in other industries, the entrepreneurs have no recourse. They must take whatever the capital will earn.

If, for example, too much specialized machinery has been constructed for the production of cotton cloth, the owners cannot transfer it to the shoe industry or any other industry, no matter what are the profits to be made in these alternative employments. Concrete examples of capital which has thus been unwisely invested are found in every community. In every city there are buildings which cannot be let to tenants at rentals which will pay both the operating expenses and the customary return on the money tied up in the building. In certain parts of the United States competing lines of railroads are too numerous; some are therefore unable to pay dividends to their stockholders.

Costs arising from interest on fixed, specialized capital and from depreciation from obsolescence and the elements are not, in the short run, opportunity costs. This same statement applies to outlays incurred by producers for raw materials purchased under long-time contracts, to the salaries of employees who must be retained to keep the organization of the business intact, and, in fact, to the price paid for any agent of production that the producer must buy or employ for a relatively long period of time. In contrast with the outlays which have been made for such agents are payments for common labor, raw materials bought as required for use, and other similar expenses. The expenses which the entrepreneur incurs for raw materials, common labor, and repairs necessary to make good the effects of wear and tear are called the prime costs of doing business. The expenses brought about by putting up buildings, constructing bridges and tunnels — that is, interest and depreciation on permanent capital — together with the salaries of higher employees, are called supplementary costs. In general, supplementary costs result from investment in durable, specialized instruments of production; prime costs result from the purchase of human services and goods that can be quickly turned into finished commodities.

In the long run it would appear that a good must sell for a price that covers both its prime costs and its supplementary costs. If it does not, entrepreneurs will cease to produce it; for even the most durable capital wears out, and if the investment in it has not proved profitable, such capital will hardly be replaced. This seems to be a conclusion from which there is no escape. Yet it is a conclusion that rests on the assumption that entrepreneurs can foresee what will happen in the future, which they cannot do. Consequently we find that, because of repeated mistakes on the part of investors and entrepreneurs, an industry may continue for a considerable period of time producing goods at costs (the sum of prime and supplementary) in excess of selling prices.

To show that it is often the most profitable procedure for producers to sell goods at something below the sum of both prime costs and supplementary costs, let us assume that the total unit cost of the good is $5, its prime cost $3, and its supplementary cost $2. If the price falls below $3, the day-to-day expenses cannot be covered and the producers will cease hiring labor and buying materials that give rise to a good which is worth less than its prime costs. But it is otherwise if the good sells for anything over $3. All prime costs are now covered and there remains over and above such costs something with which to meet interest charges and depreciation on the fixed, specialized capital. If the entrepreneur ceases producing, most of these capital expenses will go on just the same as if his plant were running at full capacity; so he will continue to produce as long as the good sells for anything above $3. But when the plant wears out, he will not rebuild it unless he anticipates that *in the future* both prime expenses and supplementary expenses can be met.

Here, however, the element of risk and the necessity for prediction and forecasting again come into play. The entrepreneur who had suffered disastrous losses because he could not cover all his supplementary costs might not be willing to risk further investments. But business conditions are not often stable, and a year or two of prosperity immediately preceding the time when the decision to invest or not invest must be made may so overshadow more remote losses that both the entrepreneur and the bond-buyer may decide again to invest in the business which in the past has not proved successful. If they decide not to invest, others may take a more optimistic view and thus keep the industry overbuilt for a long period of time. Evidently the railroads in the territory west of the Mississippi River were for years overbuilt, especially during the two decades preceding 1900. Yet this did not deter two of them from constructing at great expense two additional lines to the Pacific coast when a few years of good profits, following the revival of business in 1898, had put the older lines on a more solid financial basis.

When producers are confronted with a decrease in demand and are therefore unable to sell the maximum output of their plants, it not infrequently happens that they attempt for a time to maintain prices by reducing output either by running only part time or allowing a part of the plants to stand idle. Usually when this happens, it would be impossible to sell the capacity output for enough to cover prime cost. But output is likely to be reduced before prices have reached so low a level as prime cost. A common reason for this procedure is that to reduce prices would spoil the market and prevent a revival of demand. The prices at which the goods are sold are supposed to have become customary; and if they were reduced, consumers might refuse to buy when at a later period they were again put back to the old figure.

Whether the conditions that compel producers to sell their products at prices below the sum of prime costs and supplementary costs are permanent or only temporary depends on a number of considerations. In the depression stage of the business cycle many — sometimes all — of the producers in a trade are forced either to sell below total costs or close down. But the situation is only temporary for a majority. With the return of prosperity demand revives and costs can again be covered. But it is not with the conditions incident to the business cycle that we are here concerned, but rather with dislocations of demand and supply growing out of overexpansion of fixed plant, major changes in demand, and advances in productive technique. When these are the causes of the inability of the producers to cover all their costs, the permanence of the difficulty depends on (1) the durability of the fixed capital, (2) the trend of demand, — whether stationary, rising, or falling, — and (3) the extent to which producers are able to predict future changes in demand. For example, if the capital in the industry is very durable and if a large part of it is new, the period of readjustment may be relatively long. It may take years for the producers to reduce the amount offered for sale to a point where marginal demand price equals cost to the marginal firm. Again, if the demand is increasing, the period will be shorter under any given situation in respect to the durability of capital than it will be if demand is stationary or falling. Finally, if the producers misjudge the trend of demand and interpret each temporary revival of demand as a change in the long-time movement of demand and on that account replace old or build new plants, the readjustment of output to demand may be long in coming.

It may well be asked whether there is any truth in the statement that prices tend to coincide with costs of production in those industries which require large amounts of fixed capital. Is it not more nearly correct to say that in these enterprises speculation plays so great a part that only prime cost affects price? Is it not true that the investment in specialized capital is so largely a gamble that no relation between supplementary cost and price can exist? To these questions two answers can be made. In the first place, the price of a good cannot remain for very long above the sum of its prime cost and supplementary cost, since competition both between businesses already established and between them and new concerns will result in such an increase of production that the gap between the cost and the price will be closed. In the second place, capital must all eventually wear out and that which has proved unprofitable will hardly be replaced unless, as has been said, conditions seem to point to prices high enough to cover the going return on new investments. But, after all, too close an equality between costs and prices must not be insisted upon, for this relationship is not one established by the forces of nature apart from the action of men themselves. The tendency for the values of freely mobile agents to agree with the values of the products to which they contribute is the result of many forces operating on the demand of consumers and on the entrepreneurs who control production. A perfect coincidence between the prices marginal entrepreneurs have paid for the agents of production and the prices of the products cannot be expected as long as unforeseen forces are constantly interfering with the calculations of producers.

There are certain cases in which the coincidence is permanently interfered with. In the trade between two nations goods do not exchange for prices that yield the agents of production the same return in both countries. It may well happen that a unit of silk produced in Japan exchanges for a unit of steel produced in the United States (comparative money prices) at such a rate that labor in the steel mills in the United States can be paid higher wages than labor of comparable skill in the production of silk in Japan. The reason for this discrepancy is the fact that the Japanese workers do not have the alternative of working in American steel mills. There is even a certain amount of immobility of the workers between different sections of the United States. Consequently the same grade of labor may not receive the same wages in the production of different commodities coming from different sections. Finally, when a unique natural agent is required for the production of a single commodity, the principle of opportunity cost is not strictly applicable. The normal price of the commodity into which this agent enters is then the sum of the prime and supplementary costs when the agents having alternative uses are applied marginally on the unique agent.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Normal price is sometimes said to be the "right price"; it is also said to be the "just price" of a good. If price is above normal, in what sense is it not "right"? in what sense, if any, is it not "just"?

2. Cost of production governs price. A merchant who computes his costs correctly will never lose money. Is this statement correct?

3. Cost governs price only in so far as it affects the amount of goods to be put upon the market in the future. Hence the fact that goods now offered for sale cost a certain amount has nothing to do with price. Is this statement correct? Is it partly correct?

4. If an industry employs a large amount of fixed, specialized capital, a small increase in business may add very greatly to its profits. Is this statement true under all conditions? If not, what are the conditions under which it would be true?

5. In the long run all supplementary costs become prime costs. Why?

6. "New commodities like the radio usually sell at a price above the normal." Do you agree? Explain.

CHAPTER XI

DIFFERENTIAL COSTS AND NORMAL PRICE

Competition and Prices. In highly organized markets the most effective method of competition is price competition. If a trader on one of the large grain exchanges wishes to dispose of his holdings, he can practically always do so by offering them at a price slightly below the figure at which his competitors are willing to sell. If he wishes to increase his holdings, he can always do so by bidding a trifle above the market. In retail markets, which are not so well organized as the grain exchanges, competition through price reduction is also one of the chief methods of attracting buyers. Price increases will likewise usually bring additional sellers. But other methods, superficially different from price competition, are also employed. More effective advertising and other forms of increased sales effort, easier credit, additional services furnished with the goods sold, — all these are common methods of drawing buyers away from rivals. But a moment's reflection shows that, after all, these other forms of competition amount to the same thing as reducing prices. A better good is offered at the same price, or a special service is sold with the physical good, or the buyer is given additional time in which to pay for the good. In short, more value is offered for the same money, which is not significantly different from offering the same value for less money.

Increased use of advertising seems to be a different form of competition, but actually it is not, in many instances. Since the market for consumers' goods is not so closely knit together as is the grain market, by personal contacts and efficient reporting of information, the individual merchant must undertake to spread the news of bargains in price and quality. A merchant who took no trouble to bring his wares to the attention of consumers would be comparable to a grain-dealer who retired to an obscure corner to utter his bids and offers. When advertising aims at informing the buyer, there can be no question that it serves to organize the market and to make prices more uniform throughout the market.

Much sales effort, including a great deal of advertising, does not aim at informing the buyer but at playing upon his emotions in order to increase his estimate of the value of the goods to be sold. The aim is not to tell the buyer what prices and qualities are but to persuade him that he cannot do without the commodity for which the seller wishes to increase the demand. To a considerable degree such sales efforts have the desired effect, but this effect is much minimized by the fact that all the competitors can make use of the various selling devices. Moreover, no amount of sales effort can permanently keep up a demand for a competitive good if it is much inferior to the products of other sellers. In conclusion we must admit that while price-cutting is still the most important method of competition, other means of acquiring and holding the patronage of buyers are also in use. It does not always happen that the seller who offers the good for the cheapest price will drive out all his rivals.

An important factor limiting the ability of the seller who cuts prices to get business away from his competitors is the lack of homogeneity of markets. In retail trade, and to a less extent in manufacturing, each individual entrepreneur has a clientele. Of the group who customarily buy from him some are extremely difficult to detach and to induce to buy from other sellers; others drift about from seller to seller, constantly on the alert to discover where they can "get the most for their money." Each seller has a market to which his rivals do not have ready access, yet they do have some access. The seller cannot keep prices too high, and he must offer some bargains if he expects to keep the goodwill of his customers.

This fact is so well known that manufacturers try very hard to make their products appear different from those of competitors, in order that the consumer may be prevented from making comparisons.

The ideal market is one in which there is competition of many sellers and many buyers, in which competition is free and combination absent, in which the competition is so homogeneous that all the sellers are in contact with all the buyers, and in which the same good is sold and bought by all. These conditions are never fulfilled completely in real life. They are closely approached in the organized markets for elementary raw materials; they are approached less closely in the market for capital goods and in the retail market for some common standardized commodities; and they are farthest from realization in the retail market, especially when the good is a luxury or a very minor item in the budget of the consumer. Although complete realization of the perfect market is not attained in many exchanges, there is a constant tendency for all markets to become better organized. The great grain exchanges are a relatively recent development. Advertising has contributed greatly to the organization of the retail market, and so have improved means of communication. The education of the consumer and the dissemination of information by magazines, books, and newspapers have also contributed to this end. Hence we may conclude that the assumption of a perfect market for any given good is not contrary to fact. It is merely an assumption which, though never realized, is more nearly true for goods in general than is any substitute assumption.

Entrepreneurs' Differential Costs and Normal Price. It is not asserted that demand and supply will fix price at any given moment at the cost of any producer. Because of the inability of entrepreneurs to control supply completely and to anticipate demand, market price deviates from normal price. Normal price does, however, correspond with cost of production. But costs are not the same for all the producers who supply a given commodity. The next step in the explanation of price-making is to determine whether normal price tends to coincide with the cost of production of the most efficient, the least efficient, or some other producer. To this question two answers can be made, depending upon the conditions which are assumed to prevail within the industry producing the good whose price it is sought to explain.

The simplest assumption is that a completely static condition prevails in the industry. Population is supposed to be always the same in numbers; the amounts of capital, labor, and entrepreneurship constant; the demands of the consumers for the various commodities changeless; the state of the arts the same; and the capacity of entrepreneurs changeless. Under these conditions market price will correspond with normal price, and normal price will correspond with the cost of production of the least efficient producer in the industry. In Fig. 14 the demand, which is constant, is such that 1000 units can be sold at $1 a unit. The equilibrium price is $1 a unit, which corresponds with the cost of production of C, the least efficient firm that can remain permanently in the industry. To prove that the price cannot be anything else under these assumptions, suppose that it were $1.10 a unit. This would allow a fourth firm, D, to enter the market with an average cost of production of $1.10 a unit. The buyers would now be asked to take 1200 units, which they could be induced to do only if price fell to 75 cents. But this price would not cover the costs of any of the producers except A. Consequently all the others would tend to put the agents of production that they control to some other use, the amount of the commodity would be decreased,

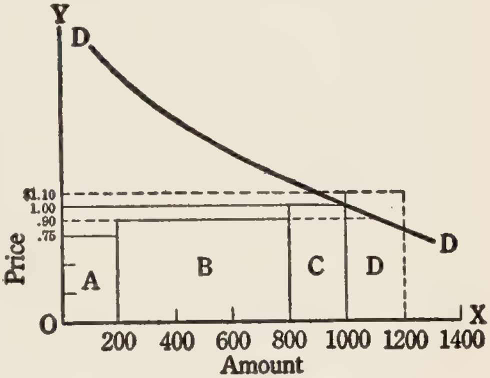

FIG. 14

and as this decrease took place, price would rise. The amount of the good might be so reduced that only 800 units were offered for sale. This would likewise result in an unstable relationship of demand and production; for that amount would sell at considerably more than $1, and producers would be able to make more in this industry than in other industries. The agents would be returned from competing industries, and price would again come to rest at $1 a unit. Evidently a producer whose costs exceeded $1 could not find a permanent place in the industry. Under static conditions, it may be concluded, normal price coincides with the cost of production of the least efficient producer whose output is required to bring about an equilibrium of demand and supply.

The conditions affecting the problem may now be changed by supposing that some of the forces considered changeless in the preceding paragraph now become variable. Let us suppose that population is increasing and that the demand for goods is subject to constant change. Under these circumstances some entrepreneurs will always make mistakes. Too large or too small a quantity of the agents of production will be found now in one industry and now in another. Because of the mistakes of entrepreneurs the agents of production will be put into industries in which their products will sell for less than they would have brought in other lines of production, or so little of these agents will be used in some industries that for a time their products will sell for more than they would in industry at large. These are the conditions of actual price-making. No industry ever adjusts itself exactly to the demand for its products for a very long period of time.

The accumulation of capital and the progress of invention, whose effects upon the volume of production are at present unmeasurable and unpredictable with any degree of accuracy, will obviously aggravate the errors due to faulty forecasting of demand. But all dynamic conditions thus far discussed tend to produce deviations of supply first on one side and then on the other side of the point of equilibrium of supply and demand. These are the conditions that prevent market price from agreeing with the normal, or stable-equilibrium, price. But all producers make errors in forecasting, and everyone expects to experience some fluctuations along with all the other competitors in the business. At times he will inevitably find his costs in excess of market prices, but at other times below. Hence we cannot say that the unpredictable nature of business inevitably causes the existence of firms whose costs are on the average above the selling price. If the diagram is taken to represent normal conditions, the unsuccessful firms cannot be accounted for by chance errors — they cannot be explained as marginal firms temporarily in distress.

Two conditions of modern business combine to give us a persistent fringe of unsuccessful producers in practically all industries. 1 In the first place, the existence of dynamic, unpredictable forces prevents business from being reduced to a mathematical routine, and therefore normal price is only the center of the fluctuations of market price. Moreover, the norm itself is constantly changing with the progress of technique and the progress of society generally. 2 In the second place, the existence of the unsuccessful fringe arises from the struggle among entrepreneurs for the monetary gains of business. All are trying to obtain the limited number of prizes offered by society for efficient production. As in any other competition some must fail, and they are the unsuccessful producers whose costs exceed price. As in any other competition the degree of success attained by different producers will be distributed from the one extreme of very low costs to the opposite extreme of very high costs. Moreover, it must be remembered that an entrepreneur may be successful for a time and later fail because of advancing age, because new conditions have arisen with which he is not trained to cope, or because he has built his business so large that it is beyond his capacity for management.

The tendency for incompetent entrepreneurs to enter business might be checked if there were an impartial jury to decide in advance whether the potential entrepreneur would be able to master not only routine but also business policy. But no such jury exists, and only by experiment can anyone determine whether he has ability to manage. We find, then, that there is a constant tendency for men to become entrepreneurs who have not the capacity to exercise that function.

Under competitive conditions the unsuccessful producers are eliminated by business failure. Because they are unable to keep their costs at or below the level of normal price, they either become bankrupt or, foreseeing that event, sell their businesses to more efficient entrepreneurs. Or they may slowly withdraw their investment by allowing their plants to wear out gradually and invest in a more hopeful enterprise, managed by others, whatever they have been able to save from the wreck. But whichever course they pursue, they are eliminated, and the agents of production which they have employed become valueless or are taken over by more efficient producers or are shifted to an enterprise where the demand for them is greater or the owner better qualified to compete.

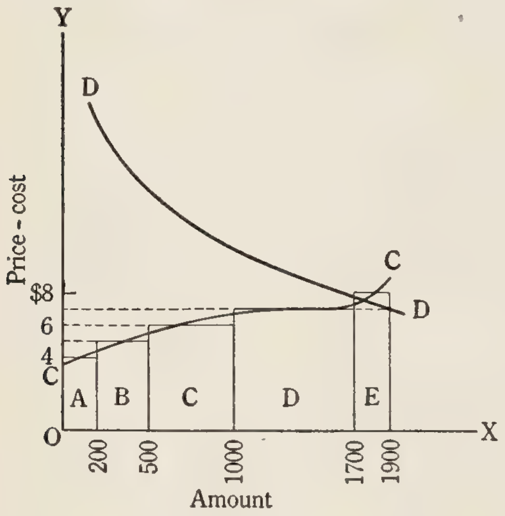

FIG. 15

But no sooner is one group of inefficient producers driven out of the business than another group appears to take its place. *In a dynamic society normal price tends, therefore, to coincide with the average cost of production of a firm or group of firms that is somewhat more efficient than the least efficient producers in the field.* This condition is illustrated by Fig. 15. The average costs of the group of five producers, *A*, *B*, *C*, etc., as represented by the upper margins of the corresponding polygons, are $4, $5, $6, etc. The curve *DD* represents the total demand for the commodity. All groups of producers except *D* and *E* have costs below the selling price, which is $7. The cost of *D* corresponds exactly with that price, and the cost of *E*, the least efficient, is $1 above the price. The bulk of the product is furnished by the marginal, or typical, firms in group *D*.[1] In any important market the number of groups who produce at different costs is far larger than the five shown by the polygons in the figure. It is more typical, therefore, to represent cost differentials by a continuous curve such as *CC*, which follows the same general trend as the irregular line formed by the upper margins of the polygons *A*, *B*, *C*, etc. All the producers whose costs are below $7 sell at a profit. The producer whose cost just corresponds to the selling price is called the marginal producer, and the supply which is just forthcoming at that cost is called the marginal supply.

Cost differentials similar to those depicted by the curve *CC* have been discovered in a number of widely different industries by actual cost studies. Of course these differentials are not the same for any two industries, but the general shape of the curve is the same. There are usually a few concerns whose costs are very low, a larger number whose costs are somewhat higher, a very much larger group whose costs nearly coincide with the selling price, and a few whose costs more or less exceed the price. As far as can be discovered at the present time, the persistence of this fringe of unsuccessful producers is a constant phenomenon in business. The proof is not, one should hasten to admit, completely conclusive, because statistical studies of costs have not been made in enough industries or under sufficiently varying conditions of business prosperity to warrant the assertion that there is no exception to the rule. There is, however, much evidence from general observation to support the conclusion. The stream of bankruptcies and failures is not intermittent or confined to a few kinds of businesses. It rises and falls with the ebb and flow of the tides of general business prosperity, but never disappears entirely.

---

[1] "Typical" is a somewhat vaguer but more realistic term than "marginal." It refers to that group of firms who achieve only moderate success, whose costs are on the average *approximately* equal to the selling price.

It would be an error to ascribe all business failures to defects in business management. Some are the result of natural forces, such as storms, floods, earthquakes, and droughts of unusual severity, and to a variety of similar disturbing phenomena. Others are the results of dishonesty of employees, of capricious legislation, and of the failure of debtors to meet their obligations. Yet the majority of bankruptcies and voluntary withdrawals are caused by lack of capital, overextension of credit, inexperience, and similar reasons that indicate incapacity of management.

Not all the firms that are found in the group having the highest costs (group *E*, Fig. 15) are on their way to bankruptcy. Some are new concerns whose managers are only getting a foothold in the business. They have business ability, but lack trade connections. In time many of them will become more efficient and find their places among the other groups that can cover their costs. Other businesses in this group are there because a single bad decision, or a year of trade depression, or a sudden expansion of supply has temporarily placed them in this position. Some of them will prove unable to win back a profitable business, whereas others will sooner or later improve their position and return to the low-cost groups. There is constant movement and interchange of position of the entrepreneurs in any industry. As a rule, however, the most efficient firms are unlikely to find themselves at any time in the *E* group; nor are the members of this group likely to move within a short period of time into the most efficient group. The interchange of position is thus limited to movement within adjacent classes.

This is the normal condition of industry. The most efficient make surpluses, or profits, while the least efficient incur losses and must either improve their position or be forced out of business. This situation appears to be unstable. The question at once arises, Why do the more efficient producers not drive all the others out of the industry? Why, when they can bring goods to market at low costs, do they not cut prices and thus push even the typical producers into the *E* group? It is, of course, to the interest of every entrepreneur to make as large an aggregate of profit as he can. But he cannot do this by selling a very small quantity of goods at a high margin of profit, nor yet by selling a large volume at a small margin. Somewhere between these two extremes a point will be found at which the product of the profit per unit of sales times the number of units sold will be at a maximum, and the volume of production represented by this number of sales is the most profitable total output for the entrepreneur to maintain. In Fig. 15, the lowest-cost producers can put 200 units on the market at a cost of $4 per unit. With the sale price at $7 their profit is $3 per unit, and their aggregate profit $600. Let us now examine the effects of an attempt on the part of these very efficient entrepreneurs to drive out the other producers. We may suppose that if they offered the good for $6 a unit, they would in time be able to sell, say, 1000 units. They would probably not be able to get all the business of their competitors immediately, because some buyers would continue to patronize other producers for personal reasons. But if the most efficient producers could sell 1000 units for $6 a unit, and if their cost of production remained $4, their pure profit would be $2000. And if they could sell 1900 units at $6, they would make a pure profit of $3800. We see, then, that if the most efficient producers can keep their costs to the low level at which they stand when these producers are furnishing only a part of the supply, they will be able to drive out all the less efficient competitors unless these competitors can also reduce their costs.

The Permanence of Entrepreneurs' Cost Differentials. The most efficient producers are not limited by inability to obtain a sufficient supply of the agents of production. Having high profits, they are able to obtain any of the agents, except entrepreneurship, by bidding a little higher than the marginal group can afford to bid. Nor is the incentive for expansion lacking, since it would appear that every extension of their output would increase profits. Yet cost differentials are found in every industry, and they do not disappear with the lapse of time. Evidently there are permanent conditions that prevent the elimination of the less efficient producers. But before an attempt is made to explain these conditions, we must recall the assumptions under which the problem is being considered. We have assumed that demand is uncertain and that producers cannot predict the amounts required to satisfy the market exactly at cost of production. But we shall rule out inventions and other discoveries tending to lower costs. As will be shown later, if the superiority of the most efficient group is derived from a cost-reducing invention, the price of the commodity will move in the direction of their cost of production.

In order to find out why the producers in group *A* do not drive out the less efficient groups, it is necessary to know what are the causes of low costs. Several explanations are commonly advanced. It is said that the possession of unusually productive natural resources, such as favorable sites or rich mines, is a prime cause of low costs. To this explanation there is, however, a very damaging objection. If the sites used by group *A* are superior to those employed by group *D*, and if this superiority is the sole cause of the low costs among the former group, the resources which are the cause of the superiority should have imputed to them a value amounting to the difference between the cost of $4 and $7. Since land is a purely physical, nonhuman element in production, other entrepreneurs could produce as much on the favorable site as do group *A*. Competition for such sites will therefore raise them in value, and their annual rentals will absorb the difference between $4 and $7; the costs of production on those sites will then be $7. It doubtless happens in many instances that the firms who use favorable sites have purchased them at an earlier time before their full value was known and that their books show only original purchase prices and not present economic values. These firms then suppose that their costs are low because, owing to an accounting convention, capital assets are seldom advanced in value. Whatever may be the justification of this convention for the purposes of accountancy, it conceals the true economic cost of production on these sites. If the land on which a factory stands can be sold in the open market for $15,000, and if land of that value bears a rental of $600, the cost of using the factory site for one year is $600, no matter what the firm may have paid for it. We must conclude that the possession of a favorable location for manufacturing, rich deposits for mining, or high-grade land for agriculture does not enable a producer to keep his economic costs below those of his competitors.

Low costs are also said to be attributable to goodwill, which enables the producer either to sell at a slightly higher price or, more probably, to retain a constant and large volume of trade. But goodwill usually has its economic cost. It is gained through superior goods, superior service, or superior advertising. If these costs are current, they will usually be counted into the expenses of production. But sometimes the goodwill is a legacy of past efficiency and past expenditures, and having been charged against the product during the time the expenditures were made, it does not show on the books of the firm at the present time. Here again, if the goodwill has a market value, it should be counted as part of the cost of production, since the firm enjoying it could sell it to another producer who could use it as efficiently as its present owners. When, however, goodwill is not vendible property but attaches to persons, the case is different. For example, a small storekeeper may have developed a valuable following of buyers in his neighborhood. He does not need to offer "specials" to hold his trade; he can select his credit risks with greater freedom and so reduce bad-debt losses; he can count on a steady volume of trade to keep his capital and his salesmen working at capacity. Part of this following has probably become attached to the store as an objective entity and could therefore be transferred to other merchants who might like to buy the store. But it is more common under the circumstances that it attaches personally to the storekeeper and cannot be transferred to anyone else. When this is the case, goodwill is a true explanation of low costs.

Patents and secret processes are sometimes listed as the causes of low costs. But like favorable sites and rich farm land, these monopoly privileges have a sale value, — a competitive-use value for which other producers will pay, — and their sale value becomes an economic cost to the user.

Again, another sort of discrepancy between accounting practice and economic analysis may give rise to a difference between costs computed from the records of the firms and the true economic cost as we have explained it. A manufacturer may find, for example, that after including in his costs all items such as rent of land, annual values of goodwill and patents, his gross sales are still above his total costs. But in this computation there is often a neglected element of opportunity cost. In many businesses it makes little difference for practical purposes whether the single entrepreneur or the partners receive the gains of the business in the form of profits or in the form of salaries. But correct economic analysis requires that the entrepreneurs should be credited with the salaries they could earn as hired managers, and these salaries likewise are a part of the economic cost of production.

The principal permanent cause of the persistence of a low-cost group side by side with several other groups whose costs are higher and with whom they are in competition is the superior management of the group whose costs are low. If after all true economic costs have been deducted, a firm engaged in a highly competitive business continues to earn a surplus, we must conclude that the entrepreneurs in this firm are able to make use of the agents of production, which they buy at the same rate as their competitors, in a superior manner. If a surplus remains after all economic costs have been deducted from the gross receipts, that surplus is pure profit; and if it is earned year after year, we conclude that the firm is superior to others because the management is superior.

Superior firms do not, as a rule, drive out all their inferiors. To some extent inertia, the tendency to "let well enough alone," is responsible for the failure of low-cost enterprises to force price so low that competitors are driven into bankruptcy or forced to shift to some other business. In other cases the firms which appear to be competing are in reality in partly different markets and on that account cannot force each other to meet price competition. But the principal reason why we find cost differentials persisting for considerable periods of time is the operation of the principle of diminishing returns. No matter how efficient a firm may be, there is a limit to the size of the organization which it can control. As it builds its business larger and larger, a point is eventually reached after which it can expand only at increasing unit cost for the additional production taken on. Finally, any added product costs as much to produce as it will sell for. In other words, the marginal cost of production in the particular business has become equal to price. At this point further expansion would result in reduced profits.[1] At this point the expansion of the superior firm ceases. It is possible that before the most efficient firm has expanded to the point where an added unit of business costs more than it sells for, it will have expelled all its competitors. Or in terms of Fig. 15, a given firm in group *A* continues to expand until it has entire control of the market and sells 1900 units. It would not be possible, however, for it to produce that amount at an average cost of $4 if its superiority rested solely on entrepreneurial efficiency. As the firm took on more business and assumed the management of a larger and larger plant, it would encounter diminishing returns from management, and in consequence its average cost of production would also increase. But it is possible that it could drive out all its competitors before its marginal cost equaled the selling price, and thus supply the entire market. If this occurred, we should have a monopoly based solely on entrepreneurial efficiency.

---

[1] The student should at this point review pages 67–78 of Chapter V.

What has been said of the most efficient producer applies as well to all competitors except those in the extra-marginal group (group *E*, Fig. 15). Each of the firms tends to push production to the point where its marginal cost and normal price agree. Hence normal price agrees neither with the average cost of production of all the competitors taken together nor with the average cost of any one of them except the marginal firms, but with the marginal cost of all, the extra-marginal group excepted. The difference between average cost and marginal cost is wide in the case of entrepreneurs of highest efficiency, lower for the less efficient competitors, and practically nonexistent for the group of producers near the margin. These marginal concerns are so poorly managed that if their true economic costs are computed (as distinguished from their book costs), they are found to agree with the normal price of the product. That is, after out-of-pocket costs have been paid and after allowance has been made for interest on proprietors' capital and for proprietors' wages, no surplus remains. The sum of these economic costs equals the normal price of the product. If the marginal firm is a corporation the stockholders, who are the entrepreneurs, receive dividends comparable with the rate of interest they could have obtained by investing their capital in securities where the risks were comparable. Entrepreneurship of their grade of ability is so plentiful among salaried employees that it has little or no economic value. In favorable years their costs are below selling prices, but in unfavorable ones they are above. Sometimes, as the result of chance interferences with competition or other favorable conditions not of their making, some of these producers make clear gains not offset by corresponding losses. But if they sometimes gain, they are at other times overtaken by chance losses, and the balance tends to be reached at zero.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Gasoline sells for approximately the same price at all the stations in a city. Does this show that there is no competition between firms? Does it show that the more efficient producers are not attempting to drive their less efficient rivals out of business?

2. In some cases the large concerns set a price which the smaller competitors follow. Does this practice show that demand and supply do not control price? Does it show that there is no competition in the trade?

√ **3.** The English economist Marshall argued that prices tend in the long run to approximate the expenses of production of the "representative firm." This firm is described as one that has normal efficiency, normal equipment, and normal conditions under which to produce and sell. Recent investigations have shown that no *one* firm remains for very long in the representative group. There is a condition of "representativeness," but the firms found in the group are constantly changing. Do these investigations overthrow the conclusion that normal price tends to agree with the expenses of the representative firm? Do they modify that doctrine?

√ **4.** The following is a list of the alleged causes of low cost among efficient firms: better management, better machinery, research laboratories, cheaper power, better location in respect to markets or sources of raw materials, lower interest on borrowed money, better or more complete utilization of raw materials, and lower loss from bad debts. Which of these causes gives rise to actually lower costs of production and which to apparent lower costs? Which causes are connected with the personal ability of the entrepreneur? Which not?

5. Suppose that a market is supplied by only three producers. Should you expect that the analysis of the relation of price and cost of marginal firms as explained in this chapter would apply? Why, or why not? Is this likely to be a truly competitive market?

CHAPTER XII

INTERDEPENDENT PRICES

The Principal Types of Interdependent Prices. Thus far in our discussion of normal price we have given little attention to the fact that the prices of two or more finished goods or two or more producers' goods may be so related to each other that a change in the demand for or the supply price of one will affect the demand for and the supply price of the other. Simple conditions have been assumed for the purpose of making clear the fundamental relations of cost and price. A few complicating conditions of the type we are about to study have, of course, been taken into account in preceding chapters. It has been shown that one consumers' good can be substituted for another if they do not differ too much in quality, serviceability, and technical use. When such a substitution is possible, 1 the price of neither good can change very much, demand remaining the same, without some effect on the price of the other. We have also shown that because many different goods are derived from the same agents of production, the normal price of one of them must be high 2 enough to attract some of the agents away from others in the same group. But these are relatively simple cases of interdependence. In this chapter some instances of more complex interdependence will be discussed.

The phenomenon of substitution is not limited to consumers' goods. Raw materials of several sorts are frequently available for the manufacture of finished goods that go by the same name and serve the same uses. A table can be constructed from a number of different woods, all so nearly alike that only the expert can distinguish them. In other cases the substitution of a different raw material makes so great a change in the appearance or the use of the finished product that two different finished goods are produced. In these cases the one raw material cannot be put in place of the other unless the consumer is willing to submit to the substitution of a different finished good. Houses can be built of brick, stone, cement, and wood, but the use of each material results in a slightly different house and a substitution of one finished good for another has occurred. Yet there is no question that the various building materials do compete with each other and that they are frequently substituted for each other in house construction. When two or more producers' goods compete among themselves for the supply of the same consumers' good, or for the supplies of slightly different consumers' goods which can still be regarded as substantially the equivalents of each other, they are said to constitute rival supplies for the finished good or goods in question. Def

The converse of *rival supply* is *joint demand.* To build a house, a contractor requires not only the materials for framing and covering, but also hardware, glass, finishing lumber, and plaster, as well as several different kinds of labor. The demand of consumers for houses is a joint demand for all these requisites for house construction plus the services of the contractor. But the term "joint demand" must not be taken too broadly. The owner of the house will doubtless require more or less furniture to make it habitable, and he will also have food cooked and served in it. But the demand for housing is not a joint demand for building materials and food or for building materials and furniture. It is likewise true that many of the things we consume in the daily round of living are used together, but this fact does not make the demand for them joint. If one of these things becomes unobtainable, we either refuse to buy the other goods or put a substitute in place of the unobtainable one if that is possible.

Joint demand and rival supply have been contrasted because the first is a condition in which the demands for the several agents are so connected that when one is wanted the others are likewise wanted. Rival supply is a condition in which the supplies of the several agents are so connected that when one is wanted (for this particular use) the others are not wanted. The distinction is shown in the following diagram, Fig. 16.

The demand for the finished good X is a joint demand for the agents A, B, and C. But it is not at the same time a demand for A, a, and a', since the latter two are rivals of agent A. In other words, A, B, and C are complements; all are required in order that the good X may be completed. But A, a, and a' are competitors in supplying X; and if the supply of A runs short, or if it goes too high in price, a and a' can be put in place of it.

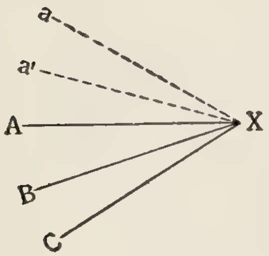

FIG. 16

When two or more different finished goods require the same agents of production, the demands of consumers for these finished goods are said to be rival demands for the agents used in producing them. The metal aluminum is used in making kitchen utensils and automobiles. Hence the demand for aluminum for making pots and pans is a rival of the demand for it in the making of automobiles. Common labor is used to build railways and to build houses, and consumers' demand for each is a rival of the demand for the other. Like joint demand, rival demand may be interpreted too broadly. All finished goods are rivals for the use of land, labor, capital, and entrepreneurship, and are therefore related in price. But this relationship is a very general one; and after specialization has been introduced, the factors divide into relatively noncompeting groups, which we have called agents of production. A true case of rivalry in demand occurs only when the same agents are wanted to produce two different finished goods. For example, the demand for cheap factory-made furniture is not a rival demand for the better wood or for the highly skilled workmanship that goes into more expensive furniture.

When one productive process or one raw material always gives rise to two or more finished goods or two or more raw materials, we have a case of *joint supply*. The classical examples are cotton fiber and seed, gas and coke, and the numerous products derived from the slaughtering of cattle. Beef is the major product of the packing house, but in order to obtain it the farmer and the packer must also produce hides, glue, tallow, fertilizer, and so forth.

Rival Supply and Price. When a finished good can be obtained from several raw materials, it may happen that only one of them will be used. Thus rails for American railroads are made of steel to the exclusion of other materials, although other kinds of iron, and wood faced with iron, have been used in the past and could be used again. Copper is likewise the only metal used in making wire for the transmission of electric current for long distances, although other metals are possible. But for technical reasons or because of expense these potential substitutes are not utilized. Hence it makes no difference what may be the price of iron-faced or iron rails; none will be substituted for steel as long as the price of steel remains near the present level. Likewise the price of platinum may rise or fall a great deal without affecting the use of copper for the transmission of electric current. At present platinum is far too costly to become a rival supply.

There are, however, many cases of rivalry in supply where the prices of the rivals are closely related to each other. Oak, birch, maple, and hard pine are used for flooring dwelling houses and other buildings. For some sorts of houses no rivalry exists, and one wood has the call of all the others. But in still other types of construction maple or birch may be so nearly the equal of oak that the builder will choose whichever is for the moment the cheaper. On still other sorts of building the rivalry may be between hard pine and maple; and if one or the other rises in price, its rival will be used.

The conditions that determine whether goods will compete with each other are (1) technical requirements, (2) tastes and whims of the consumer, and (3) the prices at which the goods can be obtained in the market. The exclusive use of steel rails and of copper wire is chiefly owing to the technically superior serviceability of these two metals for their respective employments. The use of silk instead of wool or cotton in the manufacture of certain articles of apparel is determined in part by technical conditions but chiefly by the fact that the users of cloth prefer silk. But steel supplanted iron in the manufacture of rails only after the Bessemer process had greatly reduced the cost of steel, and the use of silk in place of other fabric materials came about after the development of silk-growing and the improvement of textile machinery had made possible a reduction in the cost of silk cloth. Other woods have become competitors with oak for flooring in regions where the latter is not available and in all parts of the United States since the increasing scarcity of oak timber has greatly advanced its price.

Substitution affects prices in three ways: in the first place,
1 when the demand for a good increases, the existence of a rival
supply may prevent its price from rising as high as it would
have risen had not the substitute been available; in the
2 second place, progress of the technique of production may
bring in a competitive good in such large amounts that the
price of the good for which it is a substitute will be greatly
reduced; in the third place, the demand for a good may be
3 increased by the discovery of new uses for it, that is, when it
becomes a substitute for other goods.

The conditions of substitution may be roughly classified as follows: (1) complete substitution, (2) potential substitution, and (3) partial substitution. When substitution is complete when a new good entirely displaces the old one, there may be no price relation whatever between the two goods. Railroads are not at present likely to be displaced by ox carts, and the price at which the latter would sell, if manufactured, has nothing to do with the price of railway equipment. On the other hand, the existence of a potential substitute acts as a check upon the rise in price of the good in use; and whenever the divergence in price between a good and its substitute becomes greater than the difference in their marginal demand prices, substitution will take place on a larger or smaller scale and the rise in the price of the good or the fall in the price of its substitute will be checked.

The effects of partial substitution are much more complicated than those of complete or potential substitution. The multiplicity of goods that satisfy the same types of wants gives the principle of substitution the greatest significance in modern consumption and consequently in the price-making process in the consumers' market. The most casual observation tells us that the desire for food, for clothing, for shelter, and for amusement and display may now be satisfied by a great variety of goods. To a considerable extent the existence of this variety only results in differences in consumption habits among individuals and among different social groups. No one person is likely to make use of all the possible means for satisfying his demand for food, clothing, and other ordinary wants; and if he takes up a new article, he is likely to drop the old one entirely. On the other hand, it is true that variety in individual consumption has been greatly increased, and the typical consumer now uses a vastly larger number of goods to gratify his wants than he did a hundred years ago.

A consumer can often make use of two or more goods for the satisfaction of the same want. If the two goods have for him exactly the same utility or desirability, — that is, if the curves representing the diminishing utility of the two exactly coincide, — then he will use the one that is lower in price. But such cases are doubtless not very common. Usually there is a preference for one or the other of the two. For example, the owner of a motor car buys oil of a superior brand, A, instead of an inferior brand, B, although he might use the latter if it were cheaper than it is and if A remained at its present price. If, now, we ask what will determine the point at which substitution of B for A will take place, the answer will be found to rest on the principle of marginal utility. Let us assume, for the sake of simplicity, that the owner uses two gallons of oil every month, or twenty-four gallons a year, no matter which brand is used. If the ratio of the marginal utility of A to B is as twelve to ten, he will buy the former as long as the ratio of the prices of the two are not more unfavorable to A than twelve to ten. If, however, he comes to believe that the better brand is something less than 20 per cent superior to the poorer, he will substitute the latter for the former. If the substitution takes place on a large scale, then the dearer (and superior) oil will be reduced (provided such reduction does not depress price below expenses of production) to the point where the two goods are again in competition.

Very often the two substitutes are not capable of being put on the market in the same amounts. The output of the superior good may be limited by natural scarcity or by higher costs of production. Here we find the following adjustment of demand and supply: The more expensive good will be used by those for whom it has a relatively high utility for all purposes, and by others for whom it has a high utility for special uses. In terms of the preceding example, if the superior oil were scarce, it would be used by those who believed it to have higher utility in proportion to its price than its rival, and by others who believed that for certain sorts of work it was superior to the cheaper oil. But it should be noted that the principle of diminishing utility would operate to check the consumption of the A brand in the case of the latter type of user. If it was consumed in large quantities, its marginal utility would decline, and finally a point would be reached where it was a matter of indifference to the consumer whether he bought A at thirty cents a quart or B at twenty-five cents. If at this point the total stock of both oils was taken off the market, then thirty cents and twenty-five cents would be the normal competitive prices of the two lubricants.

To show the relation of the principle of substitution to the principle of cost of production, we may assume that the better oil cannot be produced at a cost of thirty cents. In that case the output would be gradually curtailed to the point where, marginal utility having risen, the prices of the two rivals would be proportional to their costs. If A cost thirty-five cents, then it would be withdrawn from those uses where its marginal utility was lowest (because of the refusal of consumers to pay that price for such uses) until a point was reached where its marginal utility was as thirty-five to twenty-five as compared with the B brand. It might be hastily concluded from this statement that demand had nothing to do with the relative prices of the two goods, since each would sell in the long run for a price approximately equal to its expense of production. But this conclusion is correct only on the condition that the two goods can be produced at constant cost; that is, that it costs the same amount of money per unit to produce a large quantity or a small quantity. If the costs of production vary as the outputs of the two brands are increased or decreased, then demand will control equally with cost of production.

Joint Demand. When two or more goods are used to produce a finished commodity, the purchase price which the consumer pays for the article must be divided between the owners of these goods. The same idea may be expressed by saying that the costs of the several agents necessary for the production of the good must be covered by its price. The principles according to which this price is shared by the major factors of production — land, labor, and capital — involve the theory of the distribution of the annual income of the nation, which is the subject of later chapters. But the theory of distribution does not include the more limited relation of the agents of production within the same industry, and it does not include the price relations between different grades of materials or somewhat different types of land used in producing the same good. These more restricted relationships are a part of the study of price-making.

The theory of joint demand can best be explained by the use of concrete examples. We wish to know what effect changes in the demand for the finished product will have upon the supply prices of the various factors or agents that are concerned in the production of the finished commodity. We wish to know what will be the effect upon the supply prices of the various agents when one of these agents is increased or decreased in cost, and finally we must inquire into the effect of an increase or decrease in the cost of one of the agents upon the price of the finished commodity. Let us take for our illustration the manufacture of a good, but not the most expensive, grade of furniture. We may assume that a given article — say, a table — costs a hundred dollars to manufacture and deliver to the consumer. At that price two thousand tables will be taken off the market. One hundred dollars is the normal price of the table. Its cost is made up as follows:

Raw material	$10
Unskilled factory labor	10
Skilled factory labor	20
Capital and management (factory)	20
Dealers' margins	40
Total cost	$100

The supply price of the table is thus seen to be made up of the supply prices of the agents that are necessary to produce it. These separate supply prices are, we may assume, determined by the competition of other industries. The consumer, by his demand, determines how much shall be paid for two thousand units of the article, but he has nothing to do with the distribution of the total price among the factors.

Now suppose that the capital and management costs in the factory rise from $20 to $30, and the total cost to $110. What will be the effect on the price at which tables will sell? What will be the effect on the prices that will be paid for the other agents of production? The demand for tables has presumably a certain degree of elasticity; and if the producers attempt to raise the selling price to $110, sales will decline. Let it be assumed that they fall to 1800. But if production is cut to 1800, some part of the agents must lie idle, or they must shift to other industries. If none of them can shift, then competition will force them all — except capital and management — to take lower prices for their services, the total supply price will again decline to $100, and 2000 units will be sold. On the other hand, if all the agents can quickly change without loss to other employments, the supply prices will remain as they were before the disturbance took place, except that capital and management will now cost $30 instead of $20, and the total normal supply price will be $110. A stable equilibrium of demand and supply is arrived at when 1800 units are sold at that price.[1] We conclude, therefore, that if one of the factors jointly demanded to produce a consumers' good increases in cost, and if all the other factors can be transferred to other employments without loss, then the normal price of the good will rise by the full amount of this increase.

If the agents of production can move to other types of production, the consumer bears the increase in the cost of any one of them, and of all of them, if it happens that all rise in price. This is the long-time, or stable-equilibrium, effect. But between the time when the cost of one agent rises and the time when the shifting of the agents has been completed price would not rise by the full amount. If the increase in cost was not foreseen by producers and preparation made for reducing output, price might remain as it was for a time. Then the agents would begin to move, and price would rise gradually to $110.

Let us now modify the foregoing assumption by supposing that one of the agents cannot move to another industry without very great loss. For definiteness, we shall assume that the skilled laborers cannot move unless they are willing to accept the wages of unskilled workers. Our problem is to explain how the increase in the cost of capital and management will be borne. As before, the producers will attempt to pass the added expense to the consumer; and, as before, the consumer will retaliate by refusing to buy as many as 2000 tables. There will be idleness in the industry, and the mobile factors — dealers, unskilled laborers, and the producers of raw material — will move to other industries. But the skilled laborers cannot move unless they are willing to take the wages of common labor. Some may be unemployed, or all may be employed only part of the time. If there is competition among them, their wages may fall to the point where they can do better by becoming common laborers in other lines. Wages will fall to this point unless, before it is reached, the sum of the supply prices of the various agents again becomes $100. At that figure all the agents could again be employed and the total output could be sold for $100 a unit. From the conditions laid down in this second assumption we may draw the second conclusion concerning joint demand. Under these conditions a rise in the price of one of the agents jointly demanded is borne by the immobile factor. Or it is borne by this factor until its price has fallen so low that it is more advantageous to shift it to another and lower use — in this case to employment at common labor.[1]

---

[1] Throughout this discussion it is assumed, for the sake of simplicity, that the cost of capital and management, per unit of output, remains the same (within limits) regardless of the volume produced.

In the two preceding paragraphs we have examined the effects of a rise in the cost of one of the agents demanded jointly with others for the production of a finished good. If the change is the reverse,— if the agent falls in price,— the reasoning is the same. But in this case we find that the conclusion depends upon the mobility of factors *outside* the industry. If the capital and management costs decline to $10, and if all the other four agents can be readily increased by additional supplies from other industries, then the price of the table will fall to $90. On the other hand, if the supply of skilled labor cannot be increased, then the workmen already in the trade will absorb the decline in the costs of the cheapened factor. Price will tend to remain at about $100, and the skilled labor will be paid $30 instead of $20. There will, of course, be a period of transition and unstable equilibrium during which the price of the table fluctuates about $100.

---

[1] In these illustrations the assumption is made, for purposes of simplicity, that the combination of the various contributing agents in the industry in question is rigidly fixed by the technical conditions of the industry; that is, we are assuming that definite quantities of skilled and unskilled labor are required for the operation of a given amount of capital equipment.

A decline in the demand for tables at first affects adversely all the agents employed in their production; but after time has been allowed for the transfer of the agents to other uses, the consumers again pay the sum of the normal supply prices of these agents. If all can be transferred, normal price remains at $100, but the volume of production is curtailed to the point where the demand absorbs it all at that price. If some of the factors are immobile, they must accept a decline in price, and the consumers purchase tables at something less than the former price of $100. Conversely, if the demand increases and if there is complete freedom of movement of factors *into* the industry, the only result of the change in demand is an increased volume of production with sales at the same price as before the change took place. If, as we assumed above, one of the factors cannot be increased, that factor profits by the rise in demand for the finished commodity.

The supplies of the agents of production available for any industry are practically always elastic *if time is allowed for movement into and out of the industry.* Even skilled laborers who are narrowly specialized can be increased in numbers if time is taken to train new workmen. Their numbers can be decreased too if lowness of their wages keeps young men out of the trade. Likewise capital, even though it is highly specialized and durable, wears out in time and is not replaced if the returns to it are less in the furniture industry than in other lines of manufacturing. Hence the practical problem involves the changes in the prices of the factors during the period when the shifting is either impossible or is only going on slowly.

There are also other complicating conditions, most of which we have ignored in the preceding analysis. In the first place, some degree of combination is likely to be found among all groups of producers, for unlimited competition is nearly as rare as ironclad monopoly. The advantage in resisting a decline in price or in claiming the benefit of an increase in demand rests, other things equal, with the group that can maintain the closest combination. A strong trade union among the skilled workers in the furniture industry might prevent a reduction of wages when the capital and management costs increased, and it might enable them to prevent competition from outside the industry when the demand for tables increased. It might readily accomplish these results if it were not opposed by a counter-combination among factory owners and if the owners had no other use to which their capital and management could be put. Some advantage might be retained by the union even if the owners were well organized, provided their organization was less effective than that of the laborers. But while the union might prevent a reduction of wages, it could not prevent a reduction of the number of workmen employed. The owners could not be compelled to employ their capital and management permanently at less than the going rates in industry at large.

A still further complication arises from the fact that in the determination of the supply prices of the agents jointly demanded bargaining power plays a part. Shrewdness, knowledge of the market conditions, reserves of money to be used in fighting counter-combinations, and many other forces have some effect in actual price-making. But neither side is likely to have a monopoly of these advantages for very long. And in the long run the normal forces of natural scarcity, real costs, and competitive alternatives will exert a controlling influence.

Rival Demand and Price. The rivalry of consumers for the goods they use is largely a simple competition; that is, while there are many competing buyers who demand any one good, their demand for it is based chiefly upon the desire to use it for the same purpose. There is, for example, only one type of use for a hat or a bushel of potatoes. On the other hand, many producers' goods are wanted for numerous distinct uses, and the more elementary and unspecialized such goods are the wider is the variety of demand for them. Common labor and raw materials of many sorts are examples of goods that have literally a multitude of rival demands.

When a good has several uses, some are usually more important than others. The least important use to which a good can profitably be put is called its marginal use. Thus, if we should investigate the various uses of copper, we should find that for certain electrical machines a small quantity is practically indispensable. This, we shall assume, is its most important use. Then follow less important uses until a final, or least important, use — say, roofing houses — is reached. The marginal use is defined as that employment of any good which would be first discontinued if the price of the good should rise. When copper is five cents a pound, a considerable quantity may be used for roofing; but if it should rise to a dollar, none would be employed for that purpose, and a new marginal use would then be established.

The price of any good for which there are rival demands tends to agree with its value in its marginal use. If copper is worth, in the opinion of the users, no more than five cents a pound for roofing, then it must sell for that figure in all the more important employments to which it can be put — for electrical machines, wiring, and the like. Now, because of the principle of diminishing returns, a good may have many marginal uses. Thus common labor may find its least important use in cleaning streets, in constructing railroads, and in making steel. In any industry there are many tasks that would be slighted if the wages of such labor rose appreciably. These are all marginal uses. *But they are not marginal in their entirety.* If wages rose, railroads would still be constructed, streets would still be cleaned, and steel would continue to be produced with the aid of much common labor. But some of the least important tasks, such as giving the right of ways of railroads a neat appearance, would be omitted; streets would be cleaned less thoroughly or less frequently; and various tasks in steel mills would either be dropped entirely or be performed less frequently and less thoroughly.

The value of any good in its marginal use sets the price for the entire market. It determines how much the market as a whole will be willing to pay for a given quantity of the good. If copper has been produced in such quantities that some of it must finally be used for roofing, then the price that builders stand ready to pay for it determines the price that other users will have to pay. On the other hand, it would be erroneous to suppose that the marginal use of any good solely determined its normal price. Whether copper will be used at all for roofing houses is determined by the quantity of it in the market and by the amounts other users are willing to take at prices that preclude its employment in place of other coverings. The supply of any good offered for a given use is fixed by the quantity produced and by the amounts demanded by users who are able to pay more than the prevailing price.

The theory of rival demands explains many price changes that at first sight appear inexplicable. The price of a good may advance without any apparent change having taken place in the conditions governing its total output. The wages of common labor used in making steel sometimes rise simply because railroads have undertaken new building or extensive reconstruction and increased the demand for such labor and created a scarcity of it in the mill towns. During the summer months many casual laborers leave the cities of the central part of the United States to work in the wheat fields. As a result laborers of this class who do not leave receive higher wages for city work. When the coming of winter sends these wheat-field workers back to the cities, urban wages for this class of work again decline. It follows that it is impossible to explain the price of any producers' good in terms of its total supply and the demand for it in a particular industry. Account must be taken of all the rival demands, and especial attention must be given to its marginal use or uses. Although it is true that the totality of demand for a good and the condition governing its supply determine its price, it is at the margin that changes in demand and in costs are most effective.

Joint Supply and Joint Cost. When two goods are the inseparable results of one productive process, it is impossible to ascertain the separate cost of either as far as that process is concerned. Thus cotton fiber and cotton seed, which are the joint products of the cultivation of the cotton plant, have no separable costs of growing. When two goods are jointly supplied, they must of course sell, in the long run, for a total price that will cover their joint costs. It follows that if the supply of one can be sold for only a relatively low price, the other must sell for a high price if the pair is to be produced at all. If the demand for one of these products either rises or falls, there will be a change in the price at which the other joint product can be sold. If, for example, the cotton from a given area of land can be sold for a hundred dollars and the seed for ten dollars, and if the sum of the two prices compensates the grower for all expenses, then a rise in the demand for cotton will cause the price of seed to fall. The increase in the demand will be certain to result in an augmentation of the cotton crop, which will increase the amount of seed as well as the amount of fiber. But the increased amount of seed cannot be sold at the same price as a smaller amount, its demand remaining the same, and its price must on that account decline. In like manner a decline in the demand for fiber would raise the price of cotton seed. The only definite principle of costs that is applicable to joint products is that the sum of the selling prices of such products must cover their joint costs.

When two or more goods are jointly produced, it is impossible for the producer to cut off or to add one product, when he wishes to do so in order to increase his gross receipts, without at the same time affecting the amount of the other product. Such situations can only proceed from natural circumstances over which man has little control. Many joint products, if not a majority of them, are therefore raw materials like the fiber and seed of cotton or half-manufactured goods like beef and hides. These commodities require further processing before they are ready for human consumption, and these processes entail further expenses of production. Hence the expenses incurred in producing consumers' goods are seldom joint in their entirety. Each of the products must be further processed after its emergence from the state of jointness in production, and each must bear its own separate expenses as well as such a share in the joint costs as will permit the producer to cover joint costs. When one of several joint products cannot be sold at any price, or when the costs of finishing it are above the price that consumers are willing to pay, it is not an economic good but merely waste. However, the discovery of new uses, the development of a new process, or the development of new demand may transform wastes into economic products. Examples are the by-products of the packing industry, cotton seed, and smaller sizes of logs in the lumber industry, all of which were at one time considered valueless.

It is evident that the existence of joint supply and joint cost complicates the problem of normal value. Goods tend to sell for the sum of their costs of production; but if the products are jointly supplied, separate costs cannot be ascertained. Here demand plays the major rôle in fixing the price. If one of the two products is greatly desired by those who have the income to buy it, its price will be high and its production may be pushed to the point where the supply of the by-products is so great that they become free goods; that is, they become free as far as the process of joint production is concerned, but not in respect to their separate costs. In railway transportation this phenomenon is met with when the traffic over a line moves predominately in one direction. Since all cars moved in one direction must return in the opposite direction, it is economical for the road to seek return traffic at any rate above the difference between the cost of hauling a train of empty cars and a loaded train.

Overhead Costs not Joint Costs. Not all costs that appear to be joint are really so. In merchandising, for example, it is often not practicable to segregate the costs of selling different kinds of goods. Specifically, all advertising carried by the store may have some effect upon the sales of all departments and of all goods. If it is successful, more people pass through the store; and although many articles are never mentioned in the advertisement, they benefit to some extent because more people are brought to the place where they are offered for sale, and they may justly be said to occasion some advertising expense. Yet it would be an impossible task to apportion the expense among them all. It would also be impossible to discover the proportions of the salaries of managers that should be charged to each type of commodity, although an arbitrary assignment is sometimes made. Such costs as the salaries of managers are usually carried into an account labeled "overhead" to indicate that it cannot be definitely apportioned.

Such overhead costs are similar to joint costs, and they affect business decisions in somewhat the same way. As long as the total sales are high enough to cover all costs, both those which can be apportioned and those which cannot, the store continues in business. But there are striking differences between overhead and true joint costs. When the demand for one of a pair of true joint products rises sharply and its production is increased, the other member of the pair must also be produced more freely. And, other things equal, it will fall in price. This is the principal test of true jointness. When it is applied to the goods sold in a store, we find that a demand for one commodity does not result in an increase in the supply of other commodities. On the contrary, if the demand for one group of articles rises, the store may even contract the space and selling force formerly used for the other groups and use them to market the more profitable goods.

Still another test can be applied to determine whether the costs of two or more products are really joint costs. If a railroad raises the rates on commodities moving in one direction, the amount of traffic will tend to decline. The number of loaded cars moved in that direction having been reduced, the number moving back unloaded will also be reduced, and the necessity of offering unusually low rates on the back haul will be less. But if a store finds it necessary to raise prices in one department, that fact will not reduce the amounts of goods offered in other departments or enable the concern to ask higher prices for them. Again, if the separate costs of one of a pair of true joint products are increased, its marginal supply price will rise, sales will decline, and less will be produced. The output of the other member of the pair will also be reduced, and it will rise in price. But if an increase takes place in the special cost of selling a given article in a store and its price rises, the resulting fall in sales will neither reduce the volume of other articles offered nor cause their prices to rise.

Summary of the Theory of Normal Value. Before proceeding to the next section of the theory of value and price, which is concerned with the long-time equilibrium of cost and price under the condition of increasing or decreasing demand, we may sum up briefly the principles of normal price thus far explained.

Under competitive conditions a good tends to sell for a price equal to the sum of its costs of production to the marginal producers. Its costs of production are determined in part by the competitive demands of consumers for the various consumers' goods for the creation of which the agents of production are used. An agent must, in the long run, earn as much when used to produce one good as it does in the production of other goods. This gives rise to opportunity cost. But some consumers' goods are more highly prized than others, and it might appear that the agents would on that account earn more when applied to such production than when utilized in making other things. The solution of this difficulty is that such highly prized goods will be produced in larger amounts than commodities of little significance to the consumer, and in consequence their utility will fall to the point where the last units of the agents used in making them will have the same imputed value as the last units used in making other things. The agents thus earn the same in all competitive uses.

Production is carried on by individual firms or by corporations and partnerships that conduct their business as individuals. Despite the fact that the agents of production tend to uniformity in price, different firms produce at different average unit costs. Since price cannot agree with a variety of costs, it becomes necessary to determine how entrepreneurs' differentials affect the theory of normal price. Under completely static conditions price tends to conform with the expenses of production of the least efficient producer whose supply is necessary to meet the demand, but under conditions of uncertainty and dynamic change it tends to agree with the expenses of a group of producers somewhat more efficient than the highest-cost producer. A member of this group may be called the marginal producer or the typical firm. Price cannot agree with the cost to the least efficient firm because competition is constantly eliminating those firms which, although they possess some capital of their own or are able to borrow from others, are lacking in ability to carry on business independently and without the direction of more capable persons.

The tendency of price to agree with cost of production must not be interpreted as meaning that whatever costs have been incurred by the marginal firms will necessarily be covered by the prices offered by consumers for finished commodities. This agreement will take place only if the entrepreneurs in that group have forecast demand and supply accurately. When errors in prediction take place in respect to capital investments which are short-lived, the readjustment is quickly made. These errors in prediction give rise to the fluctuations of market price about normal price. But when too much capital has been invested in specialized, durable equipment, the adjustment may be so difficult to make that for long periods of time only the prime costs and part of the supplementary costs which arise from the investment in specialized, durable equipment can be recovered by the entrepreneurs in the marginal group. Since there are no methods available at present for forecasting the conditions of demand and supply in the distant future, any investment in very durable capital is more or less of a gamble. This kind of capital may become for considerable periods a unique good, similar to special kinds of land, whose value is determined not by its cost but by the demand for the finished product and the costs of the other agents demanded jointly with it to make the finished product.

The existence of joint demand, rival supply, and rival demand does not require any modification of the doctrine that price tends in the long run to agree with the cost of production of the marginal firms. But the condition of joint supply does require such a modification. When costs are actually joint, the combined prices of the joint products must be high enough to cover the joint cost plus the separate expenses of each; but no allocation of the joint cost to each product is possible, and therefore the price of a joint product is not governed directly by cost.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "All the goods in the world constitute rival supplies for the satisfaction of the consumer's wants. Hence the prices of all goods are related to each other." Why, or why not?

2. If the demand for beef should increase greatly, how might the cost of shoes be affected? Why?

3. There are several different materials for roofing dwelling houses, some of which cost about the same, while others are much more expensive. How do you account for this? Why is each material not sold at a price considerably different from the others?

4. It is sometimes said that the transportation of freight and passengers is an example of joint supply. Is this correct? Why, or why not?

5. How, if at all, are the prices of automobiles, gasoline, and cement related?

6. A mechanic who earns eight dollars a day at his trade carves toy boats in his spare time at the rate of two a day. Is it correct to say that the labor cost of the boats is four dollars apiece?

CHAPTER XIII

CONSTANT COST

Immediate Effects of a Change in Demand. In our study of market price we found that a rise in demand is accompanied by an upward tendency in market price, and a fall in demand by a downward tendency. The volume of goods for sale over a brief period of time is usually not readily alterable, and if a shortage is to be avoided, price must advance with demand. When demand falls, price must come down to clear the market.

But these conclusions are general; they are modified, though not negatived, by special conditions. If the producers have foreseen the change in demand and have made provision for it, either by increasing stocks in anticipation of a rise in demand or by decreasing them in anticipation of a fall, the market can be cleared without a change in price. Clothing merchants anticipate the change in the demand of the public for seasonal goods, and no change takes place when the volume of purchases rises with the beginning of the new season or falls with the waning of that season and the beginning of the next. If the seller's business is one in respect to which goodwill plays a part, he often declines to take immediate advantage of rising demand, preferring to hold his trade rather than procure the largest immediate profit. Moreover, if the rise in the demand for the good is caused by a fad, the seller must proceed carefully. Consumers may want the good very much at the time, but a slight rise in price may cause them to consider its lack of fundamental utility and to desert it for a new fad. But when all these modifying conditions applicable in special cases have been considered, the fact remains that unforeseen changes in demand cause price changes.

Effects of Changes in Demand for Longer Periods. After the producers have had some time to make adjustments of supply to an increased demand, prices tend to be either the same as before the alteration of demand took place or to be higher. In the discussion of prime cost and supplementary cost we found that the use of much durable, specialized capital in the production of any good increases the difficulty of the adjustments of supply to demand. If the buyers fail to take as much as was expected when the capital was invested, the producers may be forced to sell below the sum of the prime and supplementary expenses for a rather long period of time. If there is unused productive capacity in the industry, — that is, if not all the supplementary charges are being covered, — then a rise in demand will not necessarily affect price until it becomes so great that the existing capital equipment cannot turn out the goods at the rate they are being consumed unless it is worked beyond capacity. According to the analysis given in Chapter V, costs rise when more than a certain amount of the variable agents is combined with a fixed plant. The law of proportionality tells us that after a certain point has been reached an attempt to force further physical production from a given amount of machinery and other equipment will result in reduced returns per unit of the variable agents. Since the prices the entrepreneur pays for the variable agents do not decrease, the variable unit cost of the output rises; and after still further applications of the variables have been made, the total unit cost (cost per unit for both the fixed equipment and the variable agents) also rises.

Large orders placed with plants already running to capacity — that is, running at the point of lowest total unit cost — cause costs of production to rise. The managers of the plants raise their prices, and consequently the dealers ask more for the finished commodity. This conclusion is arrived at on the assumption that the variable agents used with the fixed plant remain constant in price. But often they cost the producer more at the same time that he is forced to work beyond the capacity of his fixed equipment. To increase output, more laborers are required; and as the force is expanded, wages tend upward, less efficient workers are hired, and overtime work, for which higher rates are necessary, is resorted to. Raw materials become scarcer or more difficult to procure in satisfactory grades, and their prices go up. Finally, obsolete plants or machines are recalled to production, and their output is also higher in cost.

When a rise in demand causes plants to run above capacity, the entrepreneurs always have the option of building them bigger and thus escaping the effects of using too many men and too much raw material for the number of machines in the plant. If the rise in demand comes slowly, or if it is easily predictable, plants do expand as demand increases, and new units are added, either by new entrepreneurs or by those already in the business. But we are here considering an unpredicted rise in demand which is sufficient to induce all efficient plants to run above capacity. If building is undertaken to meet such a situation, it will often be hurried and expensive. Moreover, piecemeal additions to old plants cause costs to rise because of the difficulty of coördinating the new units with the old and because of the interruption of the operation of the old plant. We conclude that when the demand for a good produced with much durable, specialized capital rises sharply and unexpectedly, price may either remain stationary or rise. It remains stationary when the increase in demand is insufficient to employ beyond capacity all the existing efficient plants; it rises when these plants are forced to produce beyond capacity and when new plant or equipment is put in by piecemeal additions and by hurried construction.

N.B.

We have thus far considered the effects of unforeseen changes in demand upon (1) price during very short periods of time — so short that new production from the factories cannot have any influence upon the volume of goods offered for sale — and (2) price during a period long enough to allow for a hasty adjustment of plant output to the new demand. We pass now to the consideration of the equilibrium price for a still longer period of time. We shall assume that changes in demand take place so slowly that all producers have ample time to adjust themselves to the altered situation, and we shall center attention on the permanent effects of such changes. We shall ignore for the present the consequences of the ebb and flow of demand and business activity that accompany the business cycle, reserving for a later chapter the discussion of these phenomena. As we turn to the consideration of these longer periods, we find that natural scarcity and the progress of technique come to play a more and more important rôle in the analysis of price and that the effects of overutilization and underutilization of durable capital disappear.

Since the conduct of an individual business is for the most part dependent for success on the prediction of short-time variations of demand and supply and the adaptation of production to these variations, we shall find that our analysis bears more directly on questions of public policy than on private business. We shall be concerned with the analysis that must precede answers to such questions as the effect of the protective tariff upon the productivity of the people as a whole, the effect of overpopulation and of emigration and immigration, and the consequences of the extension of large-scale production.

Constant Cost. When an increase in the demand for a commodity is met by an increase of supply which takes place under conditions substantially the same as those prevailing before the increase in demand took place, the good is said to be produced at constant cost; that is, the added demand is met by added production which comes forward without appreciable change in the expenses of production incurred by the marginal producers. In Fig. 17 the long-time trend of cost is illustrated by the curve *CC*, which is a straight line running parallel to the axis *OX*. At the intersection of the demand curve *DD* and the cost curve *CC* the price will be fixed at *NP*.

The number of units put on the market will be *ON*, which is exactly the amount that the buyers will take at the price *NP*. When the demand curve moves to the position represented by *D′D′*, the price will be *N′P′*, which is equal to *NP*, and the change in demand produces no alteration of the normal price.

Conditions under which Constant Cost Prevails. A moment's reflection should cause one to doubt whether there is any commodity whose cost could remain the same if the demand should increase indefinitely. Eventually, we should suppose, raw materials would become so scarce that cost would certainly be higher. Even if the raw materials were derived from ordinary farming operations and not from mines or forests, we should suppose that the operation of the principle of diminishing returns would sooner or later cause expenses of production to rise. On the other hand, if the demand for the same goods should greatly decrease, we should suppose that the foregoing conditions would be reversed, after the decline had gone far enough, and the costs of production would consequently decline. Finally, we know that either an increase or a decrease in demand may alter the degree of specialization or division of labor prevailing within an industry. Should this happen, costs would likewise be altered.

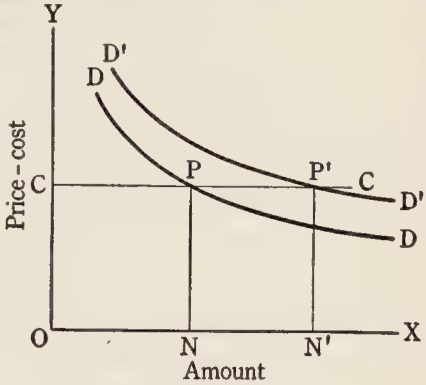

FIG. 17

It is evident that there is no commodity whose supply could be increased or decreased indefinitely at constant cost. Under certain conditions, however, the supply of any commodity may be varied considerably without any appreciable change in cost. The principal conditions are as follows: In the first place, the raw materials used in production must be so plentiful that a considerable increase in their utilization will not cause an appreciably greater scarcity. This is sometimes the case. During the period when the hardwood forests were being cleared from the land in the north-central states timber was probably available at approximately constant cost. The farmers wanted the cleared land for crops, the supply of labor was very elastic, and the forests were everywhere. It must not be supposed, of course, that no possible demand could have driven the lumbermen to production at higher costs; for if the demand had been great enough, poorer timber would have been logged and more expensive labor would have been employed. As time went on, these conditions were encountered, and the costs did go up. In the second place, as has just been 2 shown, the labor required for additional production must be forthcoming at constant cost per unit of output. This will usually be the case when the amount necessary to meet the increased demand is either relatively small as compared with the total supply of such labor for all its uses or when this labor is such that no special ability is required. Under these circumstances labor will move freely into or out of the industry without much change in wages. Thirdly, the industry must be one in which there is not much division of labor and in which the most efficient unit of production is a relatively small one. If much specialization is possible, costs will vary more or less with the increase or decrease in the total output of the industry; for with a larger output some additional specialization will usually take place and costs will fall.

One additional condition must be fulfilled if an industry is to operate under constant cost. If with increasing demand producers can introduce improved methods of using men or materials, or improved machinery already developed in other industries, costs will obviously fall. Hence, in order to complete the conditions of constant cost we must suppose that no improvements in production will accompany the expansion of the industry. But since improvements are most common in industries where the possibilities of specialization are present, this condition is practically fulfilled when we assume that the possibilities of specialization are very slight.

These conditions may be present in any industry when fluctuations in demand are not too great. Moreover, substantially the same effect will be observed if some of the agents necessary to increase the output can be secured at lower costs while others must be obtained at higher costs. Constant cost may thus result from an exact counterbalancing of decreasing and increasing costs. But such a balancing of forces cannot be expected to continue if fluctuations in demand are very great. Great changes in the demand for a commodity are likely to be accompanied by changes in its marginal cost of production.

Demand and Price under Conditions of Constant Cost. In general, price is the result of the interaction of the bids of buyers and the offerings of sellers. There are two conditions, however, under which price appears to be determined by either demand or supply alone. As we have seen, if the amount of the good for sale is absolutely fixed, demand alone will determine the price. It is also evident that if the supply of any commodity is brought to market under conditions of constant cost, the normal price will be determined by cost alone. What the demand does determine is the amount of the good that under conditions of free competition will be offered for sale.

Constant Cost and the Shifting of Costs. In the discussion of interdependent prices it has been shown that a given increase in the cost of one of the agents used in producing a commodity may cause an equal rise in its price, may cause a decline of the price of one or more of the other agents, or may cause the price of the good to rise somewhat and the price of the other agents to decline somewhat. If constant cost prevails, an increase in expenses will affect only the price of the finished article. Since constant cost occurs only when the factors of production have numerous alternative uses, it must be supposed that an increase in the cost of obtaining one of these agents is general throughout all industry, including the one in which constant cost prevails. Hence the price of the product of that industry must rise by the full amount of the increased cost if it is to be produced at all. Were the result otherwise, the agents would be completely withdrawn and the industry would be discontinued. Referring again to Fig. 17, we find that an increase in the cost of any of the agents would cause an upward shift of the line CC, throughout its entire length, to a new position above that shown in the figure. Since, as we have seen, demand determines only the amount to be produced, and not the price, the new cost curve would intersect the curves DD and $D'D'$ at a point higher on the Y-scale than OC. The opposite of this proposition is also true: a reduction of cost is followed by a fall in the price, regardless of the position of the demand curve. Moreover, when costs rise, price must rise (or fall when costs fall) by the full amount of the change in cost. Any increase in the expense of supplying one of the agents of production will be reflected by a rise in price of an equal amount; and when there is a decrease in expense, the price will fall by an equal amount.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Suppose you found that from 1900 to 1913 the cost of commodity X, when produced in typical factories, did not vary appreciably. Would this prove that the commodity was produced at constant cost? Do you need to know some additional facts before you can answer this question? If so, what facts?

2. Suppose that only 80 per cent of the productive capacity of the factories producing a good have been in use. An increase in demand now enables these factories to run at full capacity. Should you expect the total unit expense of the product to increase, decrease, or remain constant? If you found that the total unit expense decreased, should you conclude that the industry was operating under decreasing cost? Might it be operating under constant cost?

3. "We know that every dollar added to the expense of producing a commodity means higher prices for the consumer by that amount." Is this a correct statement? Might it be correct under certain conditions?

4. It is sometimes said that goods whose principal expense is common labor are likely to be produced at constant cost. Why, or why not?

5. "No good is produced at constant cost. If demand falls off, the producers will cut down production, employ fewer laborers, and buy less raw material. As a result, wages and the price of raw material will fall. If demand rises, higher prices will be paid for labor and raw material. Costs are thus shown to rise when demand increases and to fall when demand declines." Criticize.

6. The prices of many branded articles sold at retail sometimes remain the same while total consumption increases greatly. Does the fact that these articles remain stable in price prove that they are produced at constant cost? Why, or why not? What other conditions might cause their prices to remain stable?

7. Does the line *CC* in Fig. 17 show the average cost of production of all the firms contributing to the supply of *ON* units? the marginal cost of all the firms at any given time, say, when *ON* units are being produced? the average cost of certain firms?

CHAPTER XIV

INCREASING COST

The Meaning of Increasing Cost. In the preceding chapter we have seen that the product of an industry can be increased at constant cost either when the agents of production can be obtained in variable amounts without change in their prices or when the tendency to decreasing returns at one step of the productive process is offset by increasing returns at another step. In this chapter we shall advance another stage in the analysis of the effect of an increase in demand and examine the conditions under which the normal cost of production increases with an increase in the demand. A good is said to be produced at increasing cost when a rise in the demand for it is accompanied by an increase in production, which increase cannot be brought to market except at a higher cost as compared with the marginal cost of the amount produced before the demand rose.

It must be remembered that we are here concerned with long-time, or normal, effects of a rise in demand. A sharp, unexpected rise in the demand for any good may cause its costs to rise temporarily even though the permanent, or normal, conditions are such that it is produced at constant or decreasing cost. It has also been shown that if an industry is overbuilt, a sudden rise in demand does not cause a corresponding rise in cost, because the resulting increase in production does not require additional plant, and the supplementary costs can then be distributed over a larger output; total unit costs may be even lower than before the change in demand took place. But this chapter is concerned with long-run tendencies. In the long run the agents of production tend to be assigned to the productive uses in which they will earn the maximum returns.

Temporary, inefficient methods of meeting a sudden increase in demand are supplanted by permanent and more efficient methods, and plants that have become out of date or have fallen into decay are scrapped or abandoned. When these adjustments have been made, the more permanent forces assert themselves and increased demand is met by an additional supply which comes to market at constant, at increased, or at decreased cost. Here we are dealing with the conditions that make it necessary for the entrepreneurs to incur increased costs in order to provide the additional output required by the increased demand.

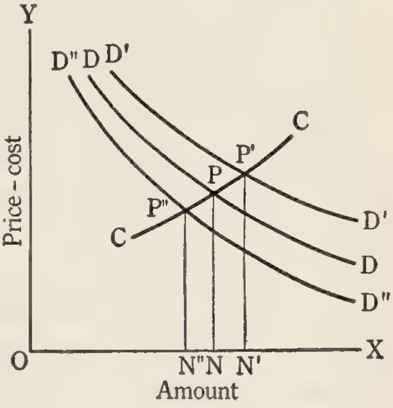

FIG. 18

The operation of the principle of increasing cost is illustrated in Fig. 18. Increasing cost is probably more typical of agricultural products than of the products of other types of production. In our explanation we shall therefore suppose that the commodity whose cost we are explaining is wheat. We also assume that before the demand increased from *DD* to *D'D'* demand and supply were in stable equilibrium with an output of *ON* and a price of *NP*. Since the equilibrium is stable, the agents of production earn as much in the production of wheat as they could earn in any other line of production.

Let us now suppose that the demand for wheat begins to increase slowly, that the demand curve moves slowly from its position *DD* to a new position *D'D'*. Producers will attempt to increase the amount offered for sale by expanding their use of the variable agents of production. Time is allowed for each producer to make the best combination he can in order to increase his product. Since the wheat is produced at increasing cost, every additional amount above *ON* will cost more than the lesser amount of *ON*. The increase in costs as amount increases from *ON* to *ON′* is indicated by the increasing distance of *CC* above *OX*. If demand becomes stationary at *D′D′*, the new equilibrium price will be *N′P′*; and since we have assumed a free competitive market, the entire output of *ON′* must sell for the price *N′P′*.

If the equilibrium of demand and production at the price *NP* is disturbed by a decline in demand to *D″D″*, costs will decrease to *N″P″*, and output will be *ON″*. It is evident that the name we have given to the phenomenon described — "increasing cost" — is not accurate, since cost *decreases* when demand falls, just as it increases when demand rises. But since the terminology is firmly fixed in economic writing, it may as well be accepted.

Conditions leading to Increasing Cost. Thus far we have been concerned with a mere description of the phenomenon of increasing cost. The illustrative figure does not explain why costs increase with rising demand and decrease with falling demand. The next step in the analysis of price is the explanation of the phenomenon just described. But before attempting to state the conditions under which a rise in demand is accompanied by increasing cost of production let us recall the fundamental conditions governing cost. In Chapter X it was explained that the costs of production incurred by the entrepreneurs in a given industry are set (1) by the competitive bidding for the agents of production by entrepreneurs making other sorts of goods, and (2) in some cases by the natural scarcity of an agent used in the given industry. But competitive bidding by all the entrepreneurs in all industries would not cause the agents to have prices (and thus cause them to "cost" something), were it not for the limitations upon their supplies arising from natural scarcity and other limitations.

When the demand for any good increases, the producers try to increase the output of the industry. To do this they must take some of the agents away from entrepreneurs producing other sorts of goods. At first they must offer slightly higher prices to overcome inertia, but in the long run they will need to pay for these agents only the same prices that other industries pay. If the increased bidding for the agents arising from the augmented demand in the industry under consideration does not force up their price, it would appear that the cost of production would not increase. But this conclusion is not correct under all conditions. It is correct (1) if the increase in the demand for the products of this industry is accompanied by a decline in the demand for some other good or goods and (2) if the agents of production which are required to make this good can *all* be drawn from the industries producing the goods for which the demand has declined. Such a transfer of demand from one industry to another should not occasion any rise in normal costs in any industry.

There is one condition, however, that may change the costs of the good for which the demand has increased. The production of this good may involve a scarce natural agent which has no competitive uses in the production of any other good. When this is the case, the agents drawn from the industries whose products are in less demand must be combined with the limited natural agent used in producing the good for which demand has increased.

But according to the principle of diminishing returns, as explained in Chapter V, the additional agents cannot be made to produce as effectively as the agents already employed. Neither can they produce as effectively as they could in the industries from which they have been withdrawn. We conclude, therefore, that the additional product can be obtained only at a higher cost.

The case we have been considering — an increase in the demand for a particular good, because of a transfer of demand — is not the only circumstance that may bring about increasing cost. Natural resources are, in any given stage of the development of technique, a limited factor of production. But population, from which are derived laborers and entrepreneurs, and capital tend constantly to increase. With the growth of these three factors production must, in the absence of changes in technique, press more and more heavily upon the land and other natural resources available in any community. Hence the principle of diminishing returns comes into operation on a much wider scale than we have considered in the preceding analysis. With the growth of population and of capital the demand for food and raw materials is increased, and men try to obtain these commodities in larger and larger amounts. But to do this they must utilize more intensively the farms, forests, and mines already being worked, and they must make use of less productive agricultural areas, less productive forests, and less productive mineral deposits. From this it follows that the added population and the added capital work under the limitations imposed by the principle of diminishing returns. In the absence of the discovery of new resources and without inventions additional food and raw materials can be obtained only at increasing costs.

It is sometimes said that raw materials and food are "increasing cost" goods. It is evident that any good must be produced at increasing cost if the growing demand for it can be met only by the use of an agent whose supply cannot be increased or if additional supplies of this agent are inferior to those already in use. But it is evident, also, that even raw materials and food can be produced without encountering the limitations of the principle of diminishing returns if new natural resources as productive as those already in use can be drawn upon as demand rises.

Diminishing Returns and Increasing Expenses of Production. In the preceding section we have explained only that as the demand for goods drawn from the earth increases, the community must, in the absence of changes in the technique of production and surplus natural resources, experience higher and higher costs of production. Higher costs have been measured in terms of units of labor and capital required for combination with the natural resources. The next step in the explanation is to show how the increased use of labor and capital per unit of output leads to increasing entrepreneurs' expense of production. In this explanation we shall assume that we are dealing with the expenses of entrepreneurs of marginal efficiency, and we shall assume that the value of money remains constant.

When individual farmers attempt to expand output, they will require additional labor and capital. But the prices they will have to pay for these additional agents are set by the competitive bidding of all entrepreneurs in all industry. Hence we may assume that the demand of the farmers for additional amounts of these agents will not appreciably affect their prices. But according to the principle of diminishing returns additional amounts of labor and capital applied to the production of additional bushels of wheat will not yield as much wheat as the earlier applications. The product per unit of the additional agents will diminish. Yet the cost of each unit of these agents to the farmers will remain the same. Hence the entrepreneurs' expense of production for the additional product will rise with every step in the expansion of the wheat crop. The *average* labor and capital expense per unit of product will not rise as much as the *additional* expense. But since the entire crop must sell at the same price, and since the added bushels cannot be sold at less than cost, the price of the entire crop will rise by the amount of the added cost of growing the marginal units.

The student will naturally inquire at this point, "What becomes of the difference between the labor-and-capital expense of growing wheat under more favorable conditions (with the 'earlier' applications of the variables labor and capital) and the marginal cost?" The answer to this question cannot be given in full until the theory of rent has been explained. But it is evident that since land is the scarce agent, — the agent which is being used past the point of diminishing returns, — the landlord will claim this difference by reason of the fact that he controls the supply of land. He will exact from the users of the land a charge called rent. This charge will equalize the difference between average and marginal cost. Each farmer will be compelled to incur in addition to his expense for labor and capital an expense to obtain the use of land. Hence the money costs of all marginal farmers (entrepreneurs) will be the same.

Entrepreneurs' Differential Expenses and Increasing Cost. In Chapter XI we have shown that entrepreneurs cannot all bring their products to market at the same average expenses of production, even when they are all working under the same external conditions. Differences in expense are caused by the varying ability of the entrepreneurs themselves. When we said that the cost of production of wheat was a certain amount, we meant that such was its cost to the marginal entrepreneur. Inefficient producers incur higher costs; and more efficient, lower costs. These differences in entrepreneurs' expenses are found in all industries. They are not due to the conditions that determine whether the product can be expanded at increasing or decreasing cost. Hence the curve *CC* in Fig. 18 must not be taken to indicate the differences in cost to different producers, which have been represented by the up-sloping curve *CC* in Fig. 15, p. 180. The curve *CC* in Fig. 18 cannot represent the varying expense of producing the different portions of the total supply of *ON*.

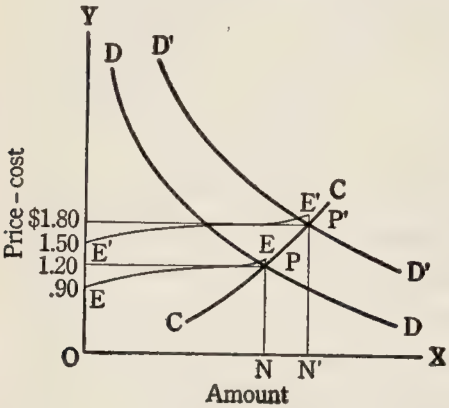

FIG. 19

In Fig. 19 the curve *CC* represents the path along which marginal cost increases or decreases as demand rises or falls. It answers the question What is the marginal cost of producing varying amounts of the good? It is identical with the curve *CC* in Fig. 18. But when *ON* units are produced, we find that entrepreneurs have average expenses of production, rent included, varying from ninety cents to more than a dollar and twenty cents a unit. For the least efficient entrepreneurs the costs are, as we know, somewhat above the selling price of the product, which corresponds with the costs of the marginal group. When demand moves from *DD* to *D′D′*, rents will rise, owing to the growing scarcity of land, and average costs will increase for all entrepreneurs. Hence the curve *EE*, which represents entrepreneurs' differential costs when demand is at *DD*, must move upward to the position *E′E′*. Every producer encounters rising costs as demand moves upward; but the more efficient are still able to keep average costs below price, the marginal group just cover expenses of production, and the least efficient still incur costs above price. As was explained in Chapter XI, all the more efficient producers will expand their outputs to the point where their marginal costs correspond with the price *N′P′*. Hence we conclude that when an industry operates under increasing cost, marginal expense and normal price rise with every permanent increase in demand and fall with every permanent decrease.

Forces counteracting Increasing Cost. Thus far we have considered the effect of increased demand upon the cost of obtaining the products of nature if there were no discoveries of new resources and if the technique of production remained constant. By "remained constant" we have implied that no inventions took place. These assumptions may now be removed and conditions brought nearer to reality. The general effect of any of these changes is to reduce the labor cost and the capital cost of obtaining raw materials. Wheat that formerly cost $1.20 a bushel may be grown for 80 cents after an improvement in farming occurs. Moreover, if the reduction in cost occurs at the same time that the demand for the product increases, it is entirely possible that the additional product might be forthcoming at $1 or even at a lower expense of production. Similarly, if the growing population discovers and cultivates new and better lands (as was the case with the surplus population of Europe during the eighteenth and nineteenth centuries) instead of cultivating the same areas more intensively, the cost of the product may actually decrease. Finally, if science reveals ways and means of tapping resources formerly unavailable, the additional population may be able to supply itself with the products of the field, the forest, and the mine without any rise of the existing normal cost.

If we were able to obtain comparable data of the cost of growing wheat or any other crop in any country over many years, we should probably find that during some of these years the cost of production was steadily rising, in others practically unchanging, and in still others steadily falling. But in every period of either falling or unchanging costs we should find that some of the conditions assumed in the statement of the law of diminishing returns had been violated and that as soon as the assumed conditions were again restored the cost of obtaining additional supplies had again begun to rise.

If the demand for raw materials and foods rises continuously, increasing cost will necessarily make its appearance unless the counteracting forces of invention and discovery are continuously accelerated. For example, we may assume that the cost of producing the marginal increments of the supply of wheat is $1.20 a bushel and that annually the growing population requires 2 per cent more than in the preceding year. If there are no improvements of any sort, the cost of getting these annual additions to supply will certainly rise. But suppose that scientists discover a new variety of wheat that either yields much more than old varieties or can be grown in areas where formerly lack of rainfall or the shortness of the growing season prevented the cultivation of wheat. The pressure of intensive cultivation on the old wheatlands will be relieved, some of the poorest will be abandoned, the marginal cost of growing wheat will fall, and the price of wheat will fall. If we assume that with every increase in the demand for wheat an invention or a scientific discovery appears and increases the productivity of existing wheatlands in the same proportion, the principle of diminishing returns will be entirely counteracted and the new demand will be met by an additional supply which the producers will be able to put on the market at a constant marginal cost.

The prices of the products of agriculture to the final consumer consist of the expenses of growing them on the farm, of processing them in factories, and of marketing. As we have seen, the first of these is sometimes less than half the price which the consumer pays for the finished good. A reduction of the costs of processing or of marketing reduces the price of the finished good, and it therefore is entirely possible that an increase in the demand for certain agricultural products might be accompanied by increasing costs of production on the farm and at the same time by reduced costs for transportation, merchandising, or processing. If these reductions are great enough, the consumers are able to secure larger amounts of the finished good at lower prices in spite of the fact that the cost of producing the raw materials has increased. For example, bread might fall in price while the cost of growing wheat increased. When the principle of increasing cost is used in the analysis of concrete problems, it is always necessary to distinguish the different steps in the complete productive process. Diminishing returns may bring about increasing costs in one process, and inventions may reduce costs in succeeding processes.

But the forces which offset the effects of diminishing returns in the production of food and raw materials can hardly be expected to keep pace exactly with the rising demand for these products. A rise in demand usually causes the price of one or more of these products to rise. If this stimulates improvements and discovery, if new resources are found or if improved methods are devised, either for agriculture or for processing and marketing, price retreats toward its original level. Very often improvements outrun the rising demand and price actually falls below the point at which it stood before demand increased. When these improvements occur in agriculture itself, the result is likely to be overproduction of farm products, and therefore many agricultural producers may find that they cannot earn the incomes they had previously earned; there follows a period of distress and discontent in the rural districts of the areas affected by the improvements.

Historically, the picture we have drawn is characteristic of many sorts of agricultural production, iron mining, coal mining, and even lumbering. Is there, then, any justification for saying that the extractive industries operate under conditions of diminishing returns? When costs are declining, is the law of diminishing returns in operation? The answer is that the tendency is always in operation. In the first place, we must note that were it not for the operation of the principle of diminishing returns, the cost under the improved conditions would have fallen even more than it did. In the second place, we must insist that the principle of diminishing returns and its consequence, increasing cost, are of the nature of universal physical laws, whose effects may for a time be obscured or counterbalanced by the ingenuity of man, but whose operation can no more be abolished or wiped out than can the law of gravitation. If the demand for wheat increases, its cost will begin to rise the moment man relaxes his efforts to perfect the arts and to delve deeper into the truths of science.

Increasing Cost and the Incidence of New Unit Costs. Under conditions of free competition, any expense necessarily incurred by the typical producers must be covered by the selling price of the goods they produce. Any increase in expense will add just that amount to normal cost. In the preceding chapter we saw that any increase in the unit expense of bringing to market a commodity produced under conditions of constant cost would raise by the same amount the normal price of that commodity. When, however, a good is produced at increasing cost, a given addition to the expense will raise the price by something less than the addition. The reason for this is evident from Fig. 20. In this figure CC represents marginal cost of production when supply is varied along OX. When ON units are produced, the cost and the price are NP. If one of the variable factors should rise in cost so that for any given amount of output the unit cost was greater than before by *PQ*, the new total cost of varying amounts of the product would be represented by *C′C′*. Producers would of course try to shift the new expense to the consumers. With the demand represented by *DD*, they could do so only by reducing output from *ON* to *ON′*. But when output has been so restricted, it is evident that the new expense is added to a smaller old expense than when *ON* units were produced, since *N′q* is less than *NP*. It is likewise apparent that the new price *N′P′* exceeds *NP* by an amount less than *PQ*.

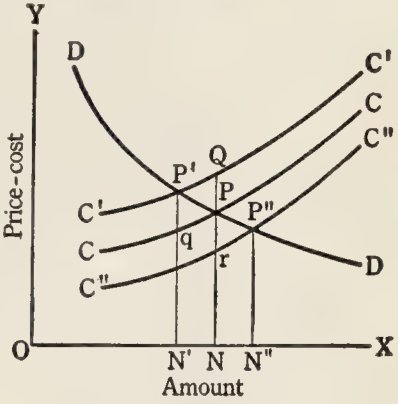

FIG. 20

The validity of this diagrammatic demonstration for real life is apparent when we recall that the producers in their efforts to shift the new cost to the consumers must reduce output. But reduction of output means the abandonment of the poorer sources of supply and the application of less of the variable factors to the fixed factor — land and other natural resources. Stated in other terms, the efforts of the entrepreneurs to shift the new expense results in the establishment of a new and lower margin of cost. The foregoing conclusion holds for taxes on the output of farms, forests, and mines. It also holds for additional labor costs and capital costs that bring about variations in the unit cost of goods. All the new cost is passed on to the consumers, but is added to a lower cost than the one prevailing before the new cost appeared.

It is sometimes supposed that any sort of subsidy paid to producers cheapens the product by the amount of the subsidy. Now a subsidy amounts to a reduction of the expense that the producers must charge the consumers in order to remain in business. Suppose that the subsidized product is being put on the market at a marginal cost of NP in terms of Fig. 20. The government now agrees to pay the producers rP for every unit sold. The effect of this reduction in expense (rP) when increasing cost rules in the industry is shown by the curve $C''C''$. Since the reduction of cost gives the producers greater profits, they will expand the output until the marginal cost is again equal to the selling price. They will increase supply from ON to ON'', and price will be $N''P''$. By the same reasoning as in the case of the added expense, we conclude that the price will fall by something less than rP. The economic argument also runs in the same terms. The producers, in order to expand the supply, must make use of poorer natural resources, and they must cultivate the old resources more intensively. Both procedures will lead to higher costs, and the subsidy will therefore be subtracted from a new and higher expense than before the subsidy was added.

In the application of the foregoing reasoning to actual cases, certain cautions must be observed. What has been said holds only if the increase in expense or the subsidy is borne or paid in accordance with some recognized unit of output. If the cost of one kind of farm machinery should rise, we might discover some — probably a very small — effect on the price of the crop in the production of which the machinery was used. But it is not necessarily true that the enhanced price would follow the principles explained in this section. Farmers might discard the machine entirely and put another in its place; they might be able to economize in its use; they might change slightly the type of product grown on the land. A further caution is necessary when we apply this reasoning to any argument concerning the effect of a subsidy to agricultural products. It has been said that such an aid would cause poorer land to be employed with a resulting increase in total product. This, in turn, would lower prices and therefore leave the farmers no better off than before the aid was granted. This conclusion may be in error in two respects. In the first place, the expansion of the farming area would be checked by increasing costs before price descended by the full amount of the subsidy, so that some benefit would still remain with the producers of all the product except the marginal increments. In the second place, the conclusion that the farmers would expand production depends upon the unstated assumption that they are already covering their expenses. If they were not, they probably would not increase their output. At any rate, they would not do so until the subsidy was made greater than the difference between expense and selling price.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. How would a curve representing increasing *average* cost (exclusive of rent) differ from *CC* in Fig. 18, which represents increasing *marginal* cost?

2. In the production of an agricultural product the following events take place: (1) the demand rises; (2) poorer land is cultivated, and the total product increases 10 per cent; (3) the cost of production on the poorer land is 15 per cent above the marginal cost existing before demand rose. Later an improvement reduces the cost of production on the better grades of land, and some but not all of the poorer land is abandoned while the product of the better land is increased. Would it be possible for the competitive price to rise at the same time that average cost stood still? If so, is the commodity produced at increasing cost?

3. When war demand increased the price of wheat, previously uncultivated areas in Montana and Canada were broken up and sowed to wheat. Do you suppose that the cost of growing wheat on these areas was higher than it was on good lands in Kansas? Why? Since 1920 the price of wheat has receded to near the level of 1913, if change in the purchasing power of money is taken into account. Yet much of the area broken up during the war is still in use. Does this show that the marginal cost of growing wheat on these lands is not higher than on the best wheatland in Kansas?

4. Suppose that all the land in a country was of uniform quality in respect to soil fertility, climatic environment, and location. As the population increased, at what point would increasing cost of the products of the land set in? Would it ever manifest itself?

5. Assume that the marginal cost of growing wheat was 90 cents in 1890, $1 in 1900, and $1.10 in 1910; assume further that the value of money remained constant throughout this period. Would a curve showing the cost of growing wheat in these years also illustrate increasing cost? ~~Expl~~ain. Why or why not ?

6. The Surplus Control Bill (McNary-Haugen Bill) of 1927, which was passed by Congress and vetoed by the President, provided for the stabilization of the prices of certain agricultural products, principally wheat, cotton, corn and swine. It was the intent of the Bill that when a "surplus" of any of these products occurred it should be taken off the market at a price determined by a Federal board and either stored or exported. It was supposed that such action would prevent overproduction from depressing price unduly.

a. Suppose that the price of any of the products should be prevented from falling to its competitive price in any year. How would the crop in the following year probably be affected?

b. Assume that the prevailing price of wheat has been $1.40 a bushel during the past five years and that the marginal cost of production has been $1.80. If the board should then raise the price to $1.80, how do you think future production would be affected?

c. Suppose the board by mistake raised the price above the competitive marginal cost for a period of years, how would future supply be affected?

d. If the surplus was sold abroad, how would foreign prices be affected? The incomes of farmers in foreign countries?

e. Do you think it possible to prevent fluctuations in the prices of agricultural products by governmental price-fixing? By storing surplus from years in which crops are abundant?

CHAPTER XV

DECREASING COST

The Meaning of Decreasing Cost. When the conditions controlling the production of any good are such that an increase in the output of the entire industry is accompanied by a decrease in the normal expenses of production incurred by the marginal group of producers, the good is said to be produced at decreasing cost. Here, as in the two preceding chapters, we are dealing with the effects of an increase in output upon the entire industry rather than upon a single plant or firm. It is often possible for a single firm or plant to expand its output at constant or decreasing total unit cost when an attempt of the entire industry to increase production would be accompanied by increasing cost at the margin of production. The expanding firm may be a new one which has not yet acquired that amount of the variable agents for combination with the fixed agents which permits the least-cost combination. In this case the management — that is, the entrepreneurship of the proprietors — is most probably the fixed factor. In agriculture it often happens that a farmer inherits, buys, or leases a farm that is too large for the amounts of machinery, live stock, and labor that he can combine with it. His land is the fixed, underutilized agent. If his management is of average quality and if he can secure capital through saving or through borrowing, he will quickly increase the variable factors, and the total unit cost of producing crops on this farm will decline for a time. It will decline until a further small increment of the variable agents applied to the fixed agent — land — results in an additional product which sells for just enough to pay the cost of the increment of the variables; that is, until the marginal costs of the principal crops grown on the farm are approximately equal to their normal prices.

But it is never possible for agricultural production as a whole to expand without a rise in marginal cost unless its expansion happens to be accompanied by a change in technique or unless the expansion takes place on newly discovered or newly opened areas. It must be assumed that the majority of the farmers have at any given time attained the combination of the agents that accompanies the lowest total unit cost under the conditions then governing agricultural production. If they had not, the desire for monetary gain would lead them all to expand their production, market price would decline, and the industry might be unprofitable for a time for a majority of those engaged in it. But when those who were least able to compete, either by reason of their low capacity as entrepreneurs or by reason of the poorness of their lands, had been driven to other occupations or into the ranks of farm laborers, price would tend to rise to a figure which just equated the supply and the demand for agricultural products. Marginal cost would agree with marginal demand price.

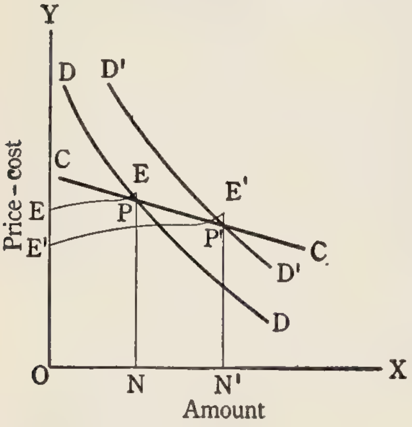

FIG. 21

The operation of the principle of decreasing cost is illustrated graphically in Fig. 21. When the total output of the industry amounts to *ON*, the cost of production for marginal producers is *NP*; and given the demand schedule represented by the curve *DD*, the normal, or long-time equilibrium, price is also *NP*. When the demand increases slowly to *D'D'*, output increases to *ON'*, and the cost of production for marginal producers declines to *N'P'*, which replaces *NP* as the normal price of the good. Conversely, when the demand declines, output recedes and cost rises. By drawing a line to connect *P* and *P′* we have the curve *CC*, which represents the path along which costs move as total production expands and contracts.

Causes of Decreasing Cost. It has already been explained that increasing cost is brought about by the operation of the law of diminishing returns; that when the conditions giving rise to diminishing returns prevail in an industry, a rise in demand is accompanied by an expansion of production at continuously rising costs of production for the entire group of firms engaged in bringing the products of that industry to market. There is a presumption, therefore, that decreasing costs will be found in those industries where increasing returns prevail and where an increase in demand is accompanied by a decline in normal costs of production for all producers.

The principle of diminishing returns, like all natural laws, is a statement of the effects following given events when certain conditions governing the events are present. When these conditions are not present, or when opposite conditions prevail, the effects are either modified or reversed. It will be recalled that the conditions under which diminishing returns accompany an increase of the output of an industry are as follows: (1) a fixed agent must be combined with increasing amounts of one or more other agents, (2) the point of best combination between the agents must have been attained, and (3) the technique of the industry must remain the same. Increasing returns can be found only when one or more of these conditions are violated.

If the factors can all be increased, so that there is no fixed factor, additional plants can be established as the demand for the product increases, and in each of these plants the combination of the factors can be as good as in the plants existing before the demand increased. If the first condition which is necessary for the operation of diminishing returns is removed, the agents of production are all variables, and all may be secured at constant prices by the entrepreneur. Since there is no fixed agent, the total output of the industry can be expanded at will until an equilibrium of price and cost is attained without affecting the marginal or the total unit costs of the various plants that make up the industry. This is essentially the condition which gives rise to constant cost.

At any given time there is a tendency for all the plants supplying a given market with a given product to approach the best combination of the agents of production. At any particular time not all the plants will actually be found in this condition, because industry is at present dynamic and change follows change with such rapidity that some of the producers fail to keep up with the advances in technique which require readjustments in the combination of the agents of production; yet in the long run all plants tend to adopt the new methods and all tend to make the best combination. *But a change in the total market demand for a commodity may affect the point of best combination of the agents of production for all the plants supplying a given market.* The classical example of effect of an increase in demand upon the combination of the agents is that of a railroad which has been built through a new and industrially undeveloped territory. The ideal combination of the agents of production cannot, in a majority of cases, be made when the road is first constructed. If there is to be any road at all, there must be a minimum investment in roadbed, stations, terminal yards, and rolling stock. Since an interchange of traffic with other roads running through dense traffic areas is usually necessary, standard sizes of equipment and a standard gauge for the roadbed are required. But the road through the new territory cannot, as a rule, find enough freight to fill long trains, or require very many trains. In either case the capital represented by the way and structures will not be utilized to capacity, and the ideal combination between this capital and other sorts of capital and labor cannot be made. Yet the best combination possible in view of the existing demand of shippers and technical requirements is made when only a few trains are operated.

Now let it be supposed that the demand for transportation increases within the new territory, because of the increase of population, the development of agriculture, and the building of factories. At first the railroad can meet this increased demand by operating longer trains. It will only need to buy additional cars and hire a few more employees. As traffic develops further, new engines as well as new cars will be needed and more laborers of all sorts will have to be employed. But no additions need be made to way and structures, although the cost of using these agents will rise somewhat, because of larger amounts of repairs when they are used more intensively. In this example the way and structures may be regarded as fixed agents of production, and the variables of rolling stock and labor can be added to these fixed agents without at first encountering diminishing returns. As traffic grows denser, diminishing returns will appear and the variable cost of moving goods will begin to rise. But because the fixed agents need not be increased, total unit cost will continue to decline even after average variable cost has begun to rise.[1] The decline of total unit cost will be checked when the rising variable cost comes to offset the decrease in the per unit expense for fixed capital. The point of lowest total unit cost will be reached sometime, and at that point it will be better economy for the road to lay down additional trackage, provided technical conditions are such that it can do so without increasing the fixed expenses too greatly. But it will not be economical to duplicate the existing fixed agents if only 10 per cent excess traffic above that which produces the least-cost combination is available.

The occurrence of decreasing cost, because of a betterment of the combination of the agents of production, can be seen very plainly in the illustration of the railroad in the new territory. But in manufacturing and marketing, although the conditions which make possible a betterment of the combination of the agents may be present, their identification is more difficult. A typical manufacturer uses many more sorts of capital goods than do railroads, and his equipment is much less standardized. The problem of making the best combination is probably more difficult, and it is also more difficult to discover the causes of variations in cost of production accompanying an expansion of the market. But the principle is the same for both types of production. Small manufacturers cannot, as a rule, maintain a condition in which every individual piece of equipment or every individual employee works with exactly the best amount of the other agents. Frequently, of course, failure to attain the best combinations among all the numerous agents must be charged against the efficiency of the management. But very often even the most successful producers find that a machine cannot be employed constantly at the sort of work for which it was designed, because the technique of machine design and of manufacture has not developed to the point where the units of equipment can be varied in size sufficiently to permit the ideally best combination to be made in all different sizes of plant or even in the most economical size under all conditions of the market.

---

[1] As was mentioned above, the repairs to way and structures will increase somewhat with more intensive utilization.

But when a manufacturing plant grows in size, it is often possible to increase some of the agents without increasing others, or without increasing them in the same proportion. What happens is that the latter type of agent is more completely employed. Again, technical experts and executives, whose work is necessary for any sort of efficient production, may not be fully occupied with the tasks they are best fitted to perform, and an increase in the size of the business may bring about a betterment of the combination of their services with other labor and with the physical equipment. Increasing returns and its counterpart, decreasing cost, may accompany the expansion of any plant when better combinations of the agents of production can be made as the size of the plant expands. When the expansion of the total market is accompanied by a betterment of the combination of the agents of production in marginal plants, the entire industry will operate at decreasing cost until the fixed agents are fully utilized.

The similarity of the effects of this particular cause of increasing returns and decreasing costs in manufacturing and in the operation of railroads involves another point. Just as it would be impossible for the manager of a railroad to go on indefinitely making better and better combinations between cars, laborers, etc., and fixed capital, such as way and structures, so it is likewise impossible for a manufacturer to go on forever making better combinations within a single plant. In both types of production the best possible combinations are eventually reached; and if the product is to be further increased, it must be done by duplicating all the agents — by building new plants.

Decreasing cost arising from a betterment of the combinations of the agents of production is of very infrequent occurrence in agriculture. This is partly due to the fact that land can be divided into whatever units the farmer pleases without loss of the contribution which natural resources make to production. Hence it is relatively easy to attain the best combination of land and other factors. Large capital investments in machinery and in power-producing equipment are uncommon, and the difficulties of adapting machines to the amount of work required are not very great. Very little of the unused productive capacity of agricultural machinery and buildings can be captured by increasing the size of the farm or by expanding the output of all agriculture. This is not to say that increasing returns from improvements in the combination of the agents of production are never encountered in the production of raw materials. In mining and lumbering the principle operates at certain steps in production.

The third condition necessary for the operation of the principle of diminishing returns is stability of the productive technique of the industry in respect to (1) inventions of labor-saving machinery and (2) scientific discoveries that reduce costs, increase the supplies of materials, or enhance the utility of products. It is admitted that if these changes in technique make limited agents more abundant, more productive, or easier to acquire, the variable factors may be increased without resulting in diminishing returns.

At this point we must distinguish sharply between two sorts of change in technique. In the first place, it is recognized that there are many labor-saving machines and processes of high efficiency in use in some industries that cannot be economically employed in other industries because the scale of operations of the latter is too small; that is, the demand for the products which the individual firms are able to reach is too small to justify the use of large and expensive machinery, the employment of experts, or the specialization of labor. In small towns and villages the mechanics who repair automobiles cannot employ all the specialized devices used in large city garages, the small power plant cannot make use of coal-crushers and automatic stokers, and the small merchant cannot employ expert buyers. But these improvements are available for the use of any plant or shop whose market is large enough to afford more or less continuous employment of them; and if the total market demand within the reach of the small concern should expand, the improved methods of the larger plant would be introduced.

An important concrete example of the possibility of adapting known improvements in a growing industry is the automobile industry during the early stages of its development. When the plants were small, many of the semiautomatic machines long in use in the production of engines and machinery could not have been applied to making cars. But it must have been apparent to all producers that these labor-saving devices would be brought into play in the automobile industry if the demand became large enough to justify the expansion of the small plants into large ones.[1]

---

[1] It must not be supposed that the decline in the cost of production of motor cars is entirely ascribed to the introduction of labor-saving devices developed in other industries, since it is well known that in many plants a large number of improvements previously unthought of, that is, inventions, have been used to cheapen production.

In contrast with the gradual installation of improved methods of production adapted from other industries are true innovations. The former type of change in technique may 2
be called adaptive improvements and the latter simply inventions. From a technical point of view the difference between the two is not very distinct, since it is probable that most improvements brought into a given industry from other industries must be modified to some extent. But from the point of view of economics the contrast is greater. It is common practice for manufacturers and merchants to take into account the savings in costs that they could make if it were possible for them to expand the market for their goods. It is also possible for those familiar with the technical requirements of an industry to predict a downward course of costs for an entire industry, because of the introduction of adaptive improvements, when they are certain that the demand for the product will grow steadily in the future. Inventions, on the other hand, cannot be predicted. Hence no producer can plan for increased production at lower cost by taking account of technical changes that are to come into being in the future.

A further distinction between the effects of adaptive improvements and inventions is brought out when we find that the latter often cheapen the cost of production in small plants and even in agriculture, where the gains from increasing the size of the plant (the farm) are usually insignificant. Inventions in the manufacture of steel and machinery which have reduced the farm cost of producing such crops as wheat did not result from an increase in demand nor were they always adaptable only on the larger farms. The fall in cost of production accompanying these improvements is attributable to the second type of change in technique — namely, invention. Very often inventions have come during a period when the world demand for farm products has been expanding, but it has not been the expanding demand which was their proximate cause. Had there been no growth of demand, they would probably have made their appearance just the same.

Historical Trend of Cost and Decreasing Cost. The curve *CC* in Fig. 21 does not represent the trend of cost of production for any commodity over a period of years. The cost of producing most of the important commodities in use today has been affected to a greater or less extent by the increased command of natural resources arising from invention and discovery. In a majority of cases the inventions which make production so much easier and life so much more comfortable than it was a century ago have been possible only because of the advancement of our knowledge of and control over the material world that has taken place during that time.

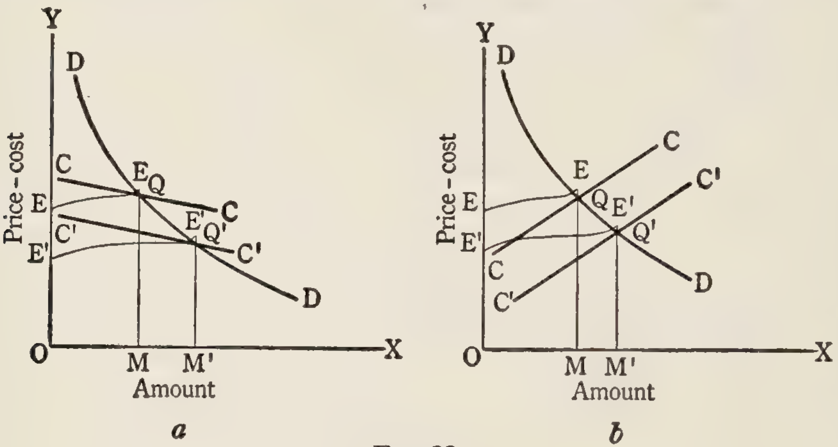

FIG. 22

The distinction between a decline in cost and normal price resulting from changes in technique which follow from an expansion of demand alone and a decline which follows from changes in technique dependent upon inventions is illustrated by a comparison of Figs. 21 and 22 *a*. In the former, demand as represented by *DD* has increased to *D'D'*, and at the same time output has increased from *ON* to *ON'*. In consequence of this enlarged market the expenses of production of the different entrepreneurs supplying the market have declined from *EE* to *E'E'*. As the demand moved upward, the entrepreneurs were able to enlarge their plants and to secure economies from adaptive improvements, from the more complete utilization of some of the agents of production, and from the greater specialization of labor. Better combinations of the agents have been made. It will usually be found under these conditions that plants have become larger and that the size of the business unit, as distinct from the size of the physical plant, has also increased.

On the other hand, Fig. 22 *a*, illustrates the consequences of an invention which reduces cost to all firms and which does not necessarily require for its adoption an increase in the size of the plants or of the business units. Demand has not increased, and yet the expenses of production of the competing entrepreneurs have declined from *EE* to *E′E′*, total production has increased from *OM* to *OM′*, and normal price has fallen from *MQ* to *M′Q′*. In this figure the curve *CC* represents true decreasing cost. If the demand should expand, entrepreneurs' expenses would decline along *CC*, since the industry is one that permits increasing returns. When invention has reduced costs, the new position of the curve representing the marginal cost of production is *C′C′* under the conditions of changed technique. But the effect of the invention is not dependent on the presence of possibilities for increasing returns in the industry. Fig. 22 *b* represents an industry in which increasing cost prevails. If an invention lowers costs from *CC* to *C′C′*, normal price will fall from *MQ* to *M′Q′*, as in 22 *a*; but at any given stage of technical development an increase in demand will lead to higher and not lower marginal cost.

In a survey of the history of industry we cannot actually separate the long-run effects of better combinations of capital and labor with natural resources from the effects of invention. The two are inextricably entangled. Large-scale production, as we know it, would scarcely have come about in the absence of inventions. But when once an invention, such as the gasoline engine, has been made, it does not follow that it will be adapted to all sorts of industry. In some productive processes the steam engine is more economical, and in others the electric motor. In still others artificial power cannot be used at all.

Whether the new source of power can be applied depends on technical requirements. It also depends on the volume of demand. In the latter case the use of the invention must await the development of the market.

Yet it is proper to make a distinction between increasing and decreasing costs as we have defined them and the actual trend of cost over a long period of time, for two reasons. First, the principles of increasing and decreasing cost are applicable to problems where the solution depends upon the combination of the factors or the agents of production. This is a problem in economics. On the other hand, a reduction of costs that results from an invention is a problem for the scientist, the engineer, or the inventor. It is primarily what is called a technical problem, and not one of economics or business.

In the second place, the effects of increasing and decreasing returns and their correlatives, decreasing and increasing costs, are principles that can be applied in purely entrepreneurial decisions. How much to produce, the combination of the factors to use, and what costs and prices will be for different amounts of output are all managerial problems. But the question of the proper machinery and the possibility of improving it by new devices are engineering problems. They become economic or business problems only after the technical stage has been passed; that is, after it has been demonstrated that the new devices can aid in performing certain productive operations. Then the question becomes one of costs and prices — whether the new machine can do the work more cheaply than it can be done by old methods.

Decreasing Cost and Restricted Competition. If the economies of large-scale production are very great, it may happen that only one firm can find a place in a market at the lowest possible unit cost. If the field is divided among several firms, the scale of operations of each does not enable it to reduce costs to the lowest possible figure. But if it happens that there are two or more competitors, either competition by price-cutting will force all but one of them into bankruptcy and leave the field to a single producer, or an agreement between the competitors to divide the market will be reached and price will be arbitrarily set at some figure above the lowest possible unit cost. None of the members of this combination will be able to produce at a cost as low as if one of them occupied the field alone. Since a combination is in existence, it is quite likely that a monopoly price will be charged consumers which is somewhat above the cost of any producer. On the other hand, it is conceivable that the price will be just high enough to enable the least efficient member of the combination to earn only his expenses of production.

But whatever the terms of the agreement, the resulting prices are likely to be unstable. Each member can increase his gains by enlarging his output if he can escape competitive price-cutting. Hence there is a tendency for each seller to seek to acquire a larger proportion of the market by surreptitiously making concessions to buyers. As any one of them enlarges his output, his costs will decline. This usually leads in the long run to an outbreak of open competition, elimination of the weaker members of the combination, and a new combination with different arrangements in regard to price. We conclude, then, that when the market will accommodate only one producer, there is no stable equilibrium price as long as more than one seller remains in the field. If one producer does succeed in gaining entire control, he will usually behave as a monopolist and raise price above cost of production.

Increasing returns do not continue indefinitely as a given firm grows in size. Whether the best size will be reached before one concern alone supplies the entire market depends on (1) the size of the market, (2) its homogeneity, and (3) the extent of the gains arising from the operation of increasing returns. The application of the first and the last conditions to the discussion in hand are easily understood from previous explanations, but the second must be considered further.

We have already seen that a market may be entirely homogeneous as far as the sellers are concerned; that is, each seller has substantially the same advantages in supplying all parts of the market area and all the buyers in that area as any other seller. On the other hand, each seller may be able to reach the entire market, but some parts will be more accessible to him than others. Each will have a partly sheltered area or a special clientele within which competition from other sellers does not operate as sharply as it does if he goes outside this advantageous zone. Furthermore, the cost of ousting competitors from their favored areas may exceed the gains that would come to any single firm from increasing returns due to enlarged output. These costs are likely to include added charges for transportation, advertising, special concessions such as more liberal credits, and minor modifications of the good sold. Hence while there is competition throughout the entire market, it is potential rather than actual in many areas. Only where the sheltered areas of different producers overlap or border on each other or in areas not yet definitely annexed to any producer will actual competition prevail.

But if the entire market expands, if the demand for the product grows, each producer's market will also expand and as a consequence costs will decline. The growth in size of the individual establishments allows adaptive improvements to be introduced and results in stimulating the producers of equipment and materials used in the industry to further specialization and reduction of cost.

Decreasing Cost and the Incidence of New Unit Costs. If an industry is operating at decreasing cost, any circumstance that measurably decreases the volume of production will force costs to a higher level. Obviously a decline in the demand for the products of the industry will also have this effect. When a new unit cost, such as a tax or a rise in the price of raw material, forces the plants to reduce output, the total unit cost will rise by more than the amount of the new cost. In Fig. 23 we have assumed that the new or added cost pushed the curve of costs from its original position of *CC* to *C′C′*. Under the original condition marginal producers could put the good on the market at a total unit expense of *NP*. But when the new cost has been added, they must ask a price that exceeds *NP* by *PQ*. Consumers will refuse to buy *ON* units at this price, the output of the industry will be curtailed, and total unit cost will rise by *rP′*, which is greater than *PQ* by *rl*. The explanation of this disproportion between the new additional cost and the rise in total unit cost is found in the fact that the producers must now give up some of the economies of large-scale production which they had previously employed.

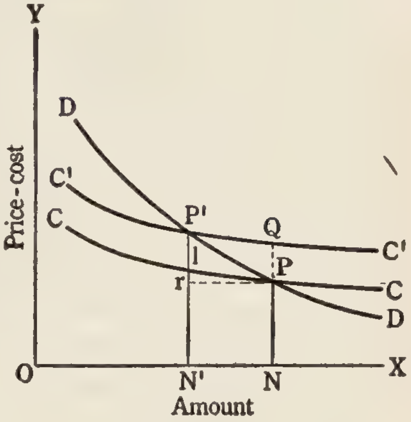

FIG. 23

Any reduction of unit expense caused by any condition other than the operation of decreasing cost itself will tend to reduce costs in an industry where decreasing cost prevails to a figure that is lower than the old costs by more than the reduction of unit expense. This is the converse of the preceding proposition. The reduction of expenses enables the producers to put the good on the market at a lower price, more is purchased, the industry expands, and costs fall still farther. By experimenting with figures similar to Fig. 23 the reader can determine just how far the costs would decline with any given assumption as to the rate at which decreasing cost takes place.

Increasing and Decreasing Cost Compared. In the absence of inventions and discoveries a growth of demand for the products of the earth is met by an ever-increasing difficulty of production. Hence there is a constant tendency for foods and raw materials to become more costly in terms of the disutilities of labor and waiting. On the other hand, the process of working up these goods into finished forms and of transporting both raw materials and finished goods has been accompanied during the past century and more by economies of large-scale production. The finished good which is thus brought to the consumer consists of utilities embodied in it as a product of nature and of utilities conferred on the raw material by men and machines in the factory, by transportation, and by the various agencies of the market organization. In the first of these operations increasing cost is the normal tendency; in the latter, decreasing cost has been the rule. It is therefore inaccurate to say that certain types of finished goods are produced at increasing cost while others are produced at decreasing cost. In the absence of inventions the raw material of the highly fabricated article just as well as the raw material of the least modified article must be produced at increasing cost. There is, however, this difference. The highly fabricated commodity owes its value chiefly to manufacture, and in manufacturing the gains from decreasing cost are likely to be prominent. Such goods tend, therefore, to fall in cost as demand for them grows. For the products of farms, mines, and forests, which come to us without much modification, the tendency, for the most part, is just the reverse.

In the case of the manufactured article it is sometimes true that the effect of increasing cost in the extractive industries is counteracted by the operation of decreasing cost in the factory. But it is significant for the solution of such public questions as the control of immigration that the two tendencies should be seen in proper relation to each other. There is no escape from the effects of the law of diminishing returns except through constant improvement in technique. Decreasing cost, as explained in this discussion, is a mere palliative. As new industries rise, we are astonished at the reduction of expense that the growth in the scale of production makes possible. Yet such industries settle down, after a longer or shorter period of time, to an even gait. When they have attained a size that permits all producers to introduce the best available mechanical devices and the highest appropriate degree of specialization, no further marked reduction in expense is possible. But if the demand for the product goes on increasing, the raw materials must be won from nature under more and more unfavorable conditions, and costs tend to rise.

The resources of the United States support at the present time a population of more than one hundred and seventeen millions. The average productivity of this population is without doubt much higher than it was fifty years ago; it is higher than it was at the beginning of this century. A part of this added productivity is due to economies of large-scale operation. A large share is also due to improvements in technique and to the discovery and utilization of some of the richest mines and the best agricultural areas in the world. These resources cannot, as far as present knowledge goes, be duplicated. Moreover, one cannot with safety predict another century of scientific discoveries such as that through which we have just passed. We can, on the other hand, predict with assurance the working of diminishing returns. Hence those who desire a further rapid increase of population in this country must either rely blindly upon future improvements to keep productivity up to its present level or they must accept as the consequence of the policy they advocate a reduction in the standard of living.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It has been contended that the process of manufacturing any good is subject to an indefinite reduction of cost, provided the demand continues to expand indefinitely so that output may likewise expand. Show that this statement cannot be defended.

2. "In 1910 industry X was very inefficiently carried on: the labor supply was poor, equipment was out of date, managers were hopelessly behind the management of other industries. In 1910 Mr. Jones entered the industry and established an up-to-date plant. Others were compelled to follow his lead, and costs fell. As a result demand for the product increased and sales doubled. Since 1910 the industry has constantly operated at decreasing cost." Is this reasoning correct?

3. "The manufacture of automobiles is a business of decreasing cost. Users of motor cars are not, therefore, affected by the fact that wages of workers in that industry have risen. Since the industry operates at decreasing cost, the managers can always find some new way of cutting other costs; the increased wages bill will be met by the industry itself." Is decreasing cost correctly defined in this statement? Is it likely that increased wages could always be met in this manner?

4. It is said that decreasing cost cannot be of very much importance in agriculture because farming operations are usually scattered over a very much greater area than is a manufacturing plant employing the same number of men. It is also said that decreasing cost cannot prevail in agriculture because the economies of large-scale production cannot be obtained. Do you agree with these statements? What is your own conclusion as to absence of decreasing cost in agriculture?

5. When an industry expands, it usually changes its methods of doing work, and it usually introduces some new machinery. Do these changes amount to a "change in technique" comparable to an invention?

CHAPTER XVI

MONOPOLY PRICE

The Meaning of Monopoly. A monopoly is said to exist whenever there is present a unified control of the supply of any commodity or service. If a single individual owns all the stock of any good, he has a monopoly of it. If he offers any part of the stock for sale, he can set any price he pleases on it without fear of underbidding by a competitor. This is a seller's monopoly. It is also possible that a buyer's monopoly may exist. As far as consumers' goods are concerned, it is quite unlikely that there would be only one person, or group of persons acting together, who could control the demand. It sometimes happens, however, that there will be only one or a few buyers for a producers' good or raw material. In these instances we may find what practically amounts to a buyer's monopoly.

The object of monopoly is usually the control of price for the benefit of the monopolist, although a governmental monopoly (the post office, for example) may exist for reasons quite apart from profit-making. In this chapter we are concerned with private or profit-seeking monopolies and their relation to the general subject of price-making. A seller's monopoly makes a profit by selling goods at prices that exceed marginal cost of production; a buyer's monopoly, by buying at prices below the competitive level. We may now complete the definition of monopoly. A monopoly exists when a person or group of persons can exercise sufficient control over either the supply or the demand for a good to enable them, under the conditions of the market, to determine its price in their own interest. A governmental board may determine the price at which a power company sells electric current, but they do not do so in their own interest and are therefore no monopoly. It may be objected to the definition that, since both demand and supply determine price, the monopolist who controls supply or demand cannot alone set a price on the article he sells or buys. This is true; and therefore we must assume that the demand or the supply is given, or, as stated in the definition, that the monopoly operates "under the conditions of the market."

In the discussion of normal value we have noted that goods do not always sell for prices exactly equal to costs of production. The discrepancies between price and cost are due partly to the inability of sellers to foresee the future relations of demand and supply and partly to hindrances to free competition. But we have also seen that these hindrances are for the most part temporary, and that in the determination of normal price they are minor factors compared with cost. Price, we have said, *tends* constantly to seek cost as a point of stable equilibrium in spite of the fact that dealers and producers frequently attempt to prevent it from falling to the cost level. In the case of complete economic monopoly, however, the supply must be so controlled that the seller can put on the market any amount of goods he pleases.

The term "partial monopoly" is sometimes used to designate those cases in which the seller or sellers do not have complete control of supply, but are able to restrict it to some extent. For example, one large producer might be able to put on the market 75 per cent of the entire output of an industry. In this case he might exercise enough influence on the supply to keep price well above the level of cost. Other producers would, of course, be entirely willing to sell at high prices. The difficulty of maintaining such a partial monopoly arises from the tendency of the minor producers to expand their output and thus cause prices to fall below the level set by the large producer; yet this tendency is frequently controlled to some extent by the fear that the large seller may cut prices below costs in an effort to ruin his competitors or to bring them back into line. It is also held in check by the inability of the smaller producers to reach the principal markets of the large seller, by their inability to secure raw materials or capital, and by mere inertia.

The Theory of Monopoly Price. If the demand for a commodity is fixed, its price depends upon the amount put on the market. It is this principle that the monopolist must use in his control of price. He cannot force buyers to pay higher prices for a *given amount* of a good than can any number of competing sellers. He must therefore limit supply in order to raise price. In Fig. 24, if *ON* units of a given commodity the demand for which is represented by *DD* are thrown on the market, they will sell for a price of *NP*. Since the cost of the supply, *ON*, as shown by *CC*, is likewise *NP*, it follows that *NP* is the competitive price and *ON* is the amount that will be offered if free competition prevails.

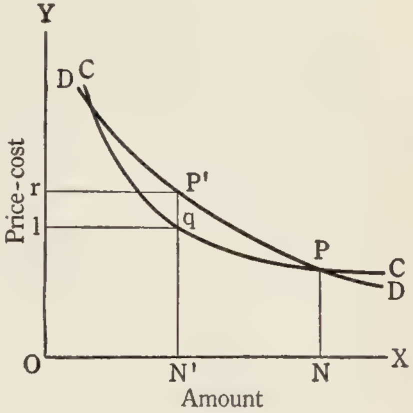

FIG. 24

Now, since the monopolist is interested in gains arising from an excess of price over costs, he will not offer as much as *ON* units for sale. How much will he offer? Clearly, since the monopoly is maintained for profit, he will put on the market that supply which will afford him the maximum excess above his cost of production. This excess is determined by two factors: (1) the amount sold and (2) the difference between cost and price. A very small amount will sell for a price that greatly exceeds cost, but the total amount of the excess — the pure monopoly revenue — will be very small. On the other hand, an amount nearly equal to *ON* can only be disposed of at a price that leaves very little monopoly gain on each unit sold. Neither a very small amount nor an amount nearly equal to *ON* would give the monopolist the maximum monopoly revenue. The point of maximum revenue is found when the product of the number of units sold by the monopoly profit per unit is a maximum. In Fig. 24 this point is reached when the amount offered for sale is *ON′* and the price is *N′P′*. The profit per unit is then *qP′*, and the total monopoly revenue, or profit, is *lrP′q*. The same principle is illustrated in the arithmetical example in Table XI.

TABLE XI

Number of Units Sold	Price per Unit	Cost per Unit	Monopoly Profit per Unit	Total Monopoly Revenue
20	$3.75	$4.00		
30	3.50	3.60		
40	3.25	3.25		
50	3.00	2.75	$0.25	$12.50
60	2.80	2.50	0.30	18.00
70	2.65	2.32	0.33	23.10
80	2.50	2.20	0.30	24.00
90	2.35	2.15	0.20	18.00
100	2.20	2.10	0.10	10.00
110	2.05	2.05		
120	1.90	2.00		

Since the good is produced at decreasing cost, a small supply cannot be economically marketed; but as the output of the industry is increased, the cost falls faster than the demand price, and when 70 units are offered, the maximum difference between price and cost per unit appears. However, this is not the amount which will yield the largest monopoly revenue, for although the monopoly profit on each unit sold is less when 80 units are supplied, the added sales more than offset the decline in profit per unit. When more than 80 units are supplied, the decline in profit per unit outweighs the increased number of units upon which the profit is realized. We find, then, that when a supply of 80 is sold at a price of $2.50 the monopolist's revenue is at a maximum.

Decreasing cost has been assumed both in Fig. 24 and in the arithmetical example. If constant or increasing cost prevailed in the monopolized industry, the point of maximum monopoly revenue would be determined in precisely the same manner; but the supplies that would yield the maximum would in all probability not be the same as if the industry operated at decreasing cost, even though the demand curves were identical in all three cases.

In the actual conduct of a monopoly the point of maximum revenue can be determined only by trial and error; for, although the variations of cost with output may be known, the market demand schedule is not ascertainable with any considerable accuracy. Moreover, demand is seldom stable through a period of time; and even though the point of maximum revenue were accurately located, it might easily be changed as new commodities came on the market or as the consumption habits of the people were altered by education, the growth of industry, and many other environmental conditions.

The Causes of Monopoly. The source of monopoly power is, as has been said, control of supply or, in rarer cases, demand. In a majority of productive enterprises the opportunity for competition is such that no single firm or group of firms can get control of the entire output. We find only a few complete monopolies in private hands. In the first place, we find that in a majority of cities only one public utility of a given type is allowed by law. Public convenience is better served by unified systems than by competing telephone, water, gas, and street-car systems. Unified systems are also likely to operate at a lower cost than are competing systems. This is true because so much of the investment of any public utility is in its trackage or in its distributive system of conduits, pipes, and wires. If competition is to operate at all, there must be at least two complete distributive systems for water, gas, and electricity or two systems of tracks serving the consumers in each area. But population is hardly ever dense enough to make possible the economical utilization of more than one system of distribution. For both reasons — greater convenience and lower costs — public utilities are usually given exclusive, or monopolistic, rights by law.

Public-utility monopolies are always subject to public regulation. They are not allowed to charge more than the price that will yield a fair return on the value of the property devoted to the use of the public. It follows that if the regulation is effective, the savings due to the unified service go to the users of the service and not to the investors in the property.

Patents and copyrights give rise to another type of legal monopoly. The inventor is given the exclusive right to manufacture, sell, and use his invention, and this right is transferable to others. Like the exclusive right to operate a public utility, the legal monopolies of invention and copyright are granted for the benefit of the public. The right of patent, it is assumed, will stimulate men to contrive improved machines, tools, and methods of doing work; the copyright monopoly is supposed to encourage authorship and scientific investigation. In both cases the public at large is the ultimate recipient of the benefits through cheaper goods, improved services, good literature, and the dissemination of scientific knowledge.

Monopolies may also be due to the exclusive ownership of scarce natural resources. Thus the De Beers company maintained for many years an almost complete monopoly of the annual new supply of diamonds by reason of its control of the mines in the Kimberley district of South Africa. It is apparent that aside from exclusive rights, such as patents and franchises, the surest method of obtaining monopoly control is the ownership of all the natural resources from which a good is derived. But the realization that this is true and the possibility of acquiring such ownership are two very different things. Sometimes a would-be monopolist supposes that he has cornered the sources of supply only to discover later that equally good or better sources in the hands of others are only awaiting development. Or having acquired the sources of supply of one commodity, he discovers that a very small rise in its price brings forth a substitute. Finally, not only are there difficulties in the way of acquiring complete control, but also the control may be rendered valueless by the fact that centralized management is inefficient. It is barely possible that all the coal mines in the United States could be brought under the ownership of one corporation, but it is doubtful if this corporation could successfully operate the mines after it had acquired them.

When an industry is operated under conditions of decreasing cost and when the gains of progressive improvement in the industry are great, there is a tendency, as we have seen, for the largest producer to drive out his competitors. When this is the case a temporary monopoly based on efficiency is theoretically possible. Yet such attempts at monopoly practically never achieve complete dominance, and when they do approach that point their tenure is usually brief. The industry eventually reaches a stage where gains from increasing size come less rapidly, and where losses from unwieldiness of the organization more than offset the gains, and competition reappears.

The possibility of profit from a monopoly of any widely used good or of any essential service is so obvious that one might suppose that all the dealers or producers would by common consent unite in a combine to control the supply of the product they sell. The reasons that such combinations are not more common or that the numerous attempts to form them are not more successful are to be found in certain human traits as well as in the conditions of the external world. Pools and other sorts of combine break down because they are unable to prevent their members from expanding the supply of the commodity they seek to monopolize. When prices rise, there are always some producers who are willing to disregard ultimate gain for the sake of temporary advantage. Outside competition can be fought by discriminatory price-cutting and other unfair methods if there are no legal restraints;

yet combinations based solely on the mere getting together of all the producers have not been permanently successful in preventing the intrusion of outsiders. The fear of legal regulation and of public opinion likewise prevents such combinations from raising prices much above the competitive level, and thus deprives them of much of the advantage to be gained from combination.

Limited Competition. In Fig. 24, *N'P'* represents the monopoly price and *NP* the price that would prevail if competition worked with faultless precision. In the preceding sections of this and other chapters we have frequently seen that a certain amount of restraint of competition might force prices above the competitive level and yet not so high as the true monopoly price. It has also been shown that the degree to which price can be forced above the competitive level will depend upon the extent to which supply can be restricted. The effect of restraints of competition on price will therefore depend upon the extent to which supply is curtailed. Since we are discussing normal price, and since we have assumed that business men are motivated by a desire for profit, we may conclude further that the effect of these restraints will depend upon the calculations made by all the producers and by potential competitors as to the most profitable use of land, labor, and capital. Let us take for our example two merchants in a small town. We may suppose that they are the only sellers who offer certain goods for sale in that market. We may further suppose that they have an agreement to sell at prices somewhat in excess of the normal competitive levels. What prices will they ask? When confronted by the question of how much effect a minor restriction will have on the price, the economist cannot give a definite answer because "a minor restriction" is itself indefinite. But the forces that would determine the action of the two sellers have all been mentioned in this and preceding chapters. The merchants would need to consider the possibility of competition from new establishments within the town, from sellers in neighboring towns, and from mail-order houses. They would need to make a nice calculation of the relative advantages to be gained from limitation of supply to a greater or less extent. Would their profits be greater if they sold 5000 units or 6000 units of goods annually? In making this calculation they would have to take account of the increased costs that might follow from a slower turnover of some sorts of goods, and of the probability that any gains from restriction of supply might be offset by later losses due to overstimulated competition. Finally, the permanence of the agreement would depend upon the possibility of such a division of the sales as to give each merchant a satisfactory share in the total profit. As has been previously pointed out, such agreements are likely to be broken by either competitor whenever he believes that the other is increasing his share of the business unduly or whenever trade is dull.

The amount of restriction on free competition that arises from such combinations is difficult to estimate. In most cases the agreement is illegal and therefore concealed as far as possible. Consumers — especially in the retail market — have very hazy notions of the competitive costs of doing business and are quick to imagine collusion among sellers where none exists. Moreover, the divergence between prices in stores offering different amounts of sales-service leads the man in the street to assume that prices may be "any figure." But economic analysis leads to a different conclusion. And the facts that business men do not all succeed, and that even in the presence of agreements many fail, give substantial backing to the conclusion that it is extremely difficult to maintain prices above the normal level unless the monopolist can get control of the supply of a good limited by nature or unless the right to produce it is covered by a patent, a copyright, or a public-utility franchise.

It follows from what has been said that the prices which may be charged by a combination that fails to achieve complete monopoly may range upward indefinitely from near the competitive level to the price that a complete monopoly would charge. For the discussion of price-making the operation of such combines differs from the operation of a true monopolist in two respects. In the first place, since the control of supply is never complete, there is always difficulty in setting a price to which all the members of the combination will agree and to which they will all later adhere. In the second place, agreements by which prices are maintained above the competitive level are likely to be short-lived, whereas a complete monopoly would continue indefinitely.

Limitations on Monopoly Price. In the preceding sections we have mentioned various reasons why a complete monopoly may not be able to derive very much profit from the control it has established; in this section we shall bring these limitations together and examine them in more detail. The first and probably most important limitation on monopoly power is the intrusion of outsiders. In the case of monopolies based on legally granted privileges or on the ownership of natural resources this limitation is absent, and such monopolies are the only ones that are likely to be at all permanent. In the second place, the consumer has a very strong defense in his power of substitution. If the producers of oranges should raise the price a great deal, some other fruit could be put in place of oranges. This is true of practically every commodity except the most vital necessities of life. For most commodities a considerable rise in price means that this price is, for many consumers, out of proportion to the marginal utility of the good expressed in terms of money, and that they will therefore cease to buy it. To be sure, the monopolist expects to cut off some of the demand, but he relies on the purchases of those to whom the marginal utility of the good is very high, or, stated in another way, he depends upon the inelasticity of the demand. But, as we have seen, substitution is a powerful factor in causing the demand for any single article to be elastic. Fortunately there are very few commodities for which there are no near substitutes. Hence to be effective a monopolist must control a good for which consumers cannot easily substitute another good, or he must extend his control over all the goods that constitute a potential supply for the satisfaction of a large group of similar human wants. The final defense of the consumer is governmental control. For centuries the common law has professed to abhor monopolies and restraints of trade, and in the United States both the national government and the state governments have set up more or less effective regulations to prevent them. It is doubtless true that these restraints do not function with the highest degree of efficiency; but there is always the possibility that the public and its officials will rouse themselves to drastic action.

PROBLEMS FOR DISCUSSION

1. Assume that a monopolist has a choice of two goods, either of which he might control. The demand for one good is elastic and for the other inelastic. Which good would offer him the better opportunity for making monopoly profit? Why?

2. It is commonly said that monopoly is more likely to occur in industries in which the principle of decreasing cost operates than in those which are governed by constant or increasing cost. Why?

3. A manufacturer requires all stores dealing in his product to sell it at a uniform, unvarying price throughout the United States. Does this uniformity show that the manufacturer has a monopoly of the good?

4. The regulations of the Federal government give a seller the exclusive right to use a distinctive trade-mark. Does this amount to a monopoly similar to a patent or a copyright?

5. If we assume that the economies of large-scale production are very great, is it to the advantage of the public to allow a few large producers to dominate the production of any commodity? under all circumstances?

6. The collective-demand curve is made up of the demands of many individuals. Those who stand ready to offer high prices for a good are nevertheless able to buy at the same price as the marginal demanders. Can you think of any devices by which the monopolist might be able

to sell to the marginal buyers at one price and at the same time exact higher prices from those whose demand prices are higher?

7. Assume three industries operating, respectively, at constant, increasing, and decreasing cost. Assume, further, that the demand for the products of the three industries is inelastic. In which industry would a monopolist prefer to establish himself? In which would he be most likely to gain a foothold? In which least likely? If you cannot answer the question as stated, what further conditions or assumptions are necessary before it can be answered?

8. Suppose that a single firm got complete control of the production of an important commodity for which there were no good substitutes, and that the monopoly was due to the entrepreneurial efficiency of that firm. Under ordinary conditions do you think that this firm would be able to charge prices very much in excess of competitive prices? Why, or why not?

9. Monopolies are detrimental to the interests of the public because they raise the prices of goods above the competitive level. To do this they must restrict production to something less than the amount that would be produced under competitive conditions. Show that such restriction may have other detrimental effects upon the productivity of the country as a whole.

CHAPTER XVII

MONOPOLY AND RESTRICTED COMPETITION

Competition tends to eliminate surplus gains, at least for marginal entrepreneurs, and constantly endangers the profits of superior entrepreneurs. If a certain field of industry is particularly bright for profit-making, its very attractiveness draws numerous entrepreneurs into the field. The more profitable the industry the more severe will be the competition. For an illustration witness the automobile industry. In the long run competition tends to reduce the attractive profits; in fact, the industry may become so overbuilt that instead of profits, heavy losses are sustained.

This situation is one of the conditions likely to lead to combination. If the industry is overbuilt, the supply must be restricted in order to sell the product at a profitable price. Combination may also be attempted, not because profits are low, but in the hope that still greater profits may be made. Combination may take a variety of forms. In this chapter we shall discuss briefly, first, various forms of combinations, second, attempts at governmental regulation, and third, price policies in an organized economy. Reas

Forms of Combination. The simplest form of combination is the pool. The pool is an agreement between independent producers with a view to limiting competition. This may be done in a variety of ways. The profits, or earnings, pool is an arrangement by which all the profits, or earnings, are turned into a common fund and divided among the various producers according to a pre-arranged plan. The earnings pool was common in the railroad field in the seventies. Thus the trunk lines from Chicago to New York pooled their through traffic earnings Pool and allocated a definite share of the earnings to each railroad. The incentive to rate-cutting was thus withdrawn. Another form of pooling arrangement is the division of territory. Each company is given exclusive privileges in a certain territory, the other companies agreeing to keep out, thus eliminating competition. Another form is the agreement to limit output, each company being assigned a certain production per annum which may not be exceeded without penalty. Sometimes a central selling bureau is established, to which all the products are turned over and through which the sales are made. The central sales bureau, together with limitation of output, is the usual form of combination in Germany, and is known as a Kartell.

In the United States pools flourished particularly in the seventies and eighties, though they have by no means been unknown since that time. Such contracts in restraint of trade and competition have never been recognized as valid in our courts. That is to say, if a member of a pool violated the agreement, he could not be sued for damages resulting to the other members because of his breach of contract, since the contract would not be sustained in court. Finally, legislation was passed, both state and Federal, which made these agreements not only invalid but even punishable by law.

trust

Another form of combination resorted to in the eighties was the historical, or technical, trust. Here we must discriminate carefully in our use of terms. Today we popularly refer to any large combination with monopoly or partial monopoly control as a "trust" without regard to its particular form. Originally the trust meant a special kind of combination. In this arrangement the controlling stockholders of the various companies entering into the combination turned their stock certificates over to a board of trustees who held the stock in trust for them as beneficiaries. Trust certificates were issued in exchange for the original stock certificates. In this manner the trustees secured control of a sufficient amount of the stock of each of the companies entering into the combination to enable them to control the various companies. They were thus able to elect the board of directors of each of the companies and so to secure a unified and harmonious management of the formerly competing concerns which had now entered into combination.

In the late eighties and early nineties trusts (in the sense here used) were declared by the courts to be illegal combinations. A different form of combination was then resorted to — the holding company. The holding company is a corporation which owns in its own name the stock of subsidiary companies. Instead of a board of trustees holding legal title to the stock of various companies, we now have a separately organized corporation owning outright the controlling stock of the different companies. The holding company therefore is able to elect the boards of directors of the subsidiary companies and thus control the policy of all.

merger

Finally, we have the merger, or consolidation. Here we have a number of separate corporations dissolving their separate identities and uniting in a single corporation which takes over the property of the several companies. Unlike the pool, trust, or holding company, the companies combining in a merger, or consolidation, lose their identity as independently existing corporations.

Besides the forms of combination proper that have been enumerated, we must also take note of less formal methods of escaping the evils (from the producers' standpoint) of competition. A common device is the interlocking of directorates, which means that one or more men serve on the boards of directors of several ostensibly competing concerns. Or even though there is no interlocking of directorates, a common policy may still be established between various companies through the indirect control by stockholders who dominate the various companies. Moreover, we have had examples, particularly in the steel industry, of informal understandings known as "gentlemen's agreements." More recently we have had "open price" associations, which aim to supply information to members about market conditions, supply and demand, orders, stocks on hand, costs and prices, but which do not directly seek to control supply or price.

The Sherman Anti-Trust Law. In 1890 Congress passed the Sherman Anti-Trust Law. The essential features of this law are contained in two sections. Section 1 states that every contract, combination, or conspiracy in restraint of trade among the several states or with foreign nations is declared illegal; section 2 declares that every person who shall monopolize or combine or conspire with persons to monopolize any part of the trade or commerce among several states or with foreign nations shall be deemed guilty of a misdemeanor. It is clear that the words "contract, combination, or conspiracy in restraint of trade" cover pools, trusts, holding companies, mergers, interlocking directorates, etc. The second section is broader. It includes not merely combinations but single corporations or individuals who seek to monopolize the market.

One of the earliest cases under the Sherman Law was the Sugar Trust case decided by the Supreme Court in 1895.[1] The government was seeking to dissolve the combination. The case involved the acquisition of a number of concerns manufacturing sugar in the state of Pennsylvania. The court held that while it might be admitted that this combination was seeking to monopolize the *manufacture* of sugar, no adequate proof had been presented in the case to show that the combination constituted a restraint of interstate *trade* or *commerce*. The combination was therefore not dissolved. It appears that the case was lost largely because the government failed to establish the direct effect of the combination on interstate trade.

Later (1899) in the Addyston Pipe and Steel Company[2] case it was decided that this combination was illegal, because the various manufacturing companies had entered into contracts which directly restrained interstate commerce. Thus, at last, manufacturing combinations were held to come within the scope of the law.

[1] 156 U.S. 1.

[2] 175 U.S. 211.

Another leading case was the Debs case,[1] decided in 1895. A strike had been called involving various railroad systems centering in Chicago. Injunctions were issued against the strikers as conspirators in restraint of interstate trade. Thus the Sherman Law was made to apply to labor unions, and the great manufacturing trusts, which Congress particularly intended to reach, at first escaped.

The "Rule of Reason." Two important railroad cases were decided in 1897 and 1898. They were the Trans-Missouri Freight Association[2] and the Joint Traffic Association[3] cases — the former in the western territory, and the latter in the eastern territory. In both cases the court held that these rate agreements were illegal under the Sherman Law. The court was divided, however. The majority of the judges held that the agreements were illegal, since the first section of the Sherman Law declared *every* contract or combination in restraint of trade to be illegal. The minority held that well-established legal usage of the terms in question would not warrant this construction. They held that under the common law it had never been the law that *every* contract in restraint of trade was ipso facto contrary to public policy. Only contracts which unreasonably restrained trade were invalid. For example, a man may sell his store under contract not to engage in the mercantile business in that city for five years. Such a contract obviously restrains trade, but it is a reasonable restraint on trade and not contrary to public policy. The minority judges held that the first section of the Sherman Law was to be interpreted in the light of the second section: the second section emphasized the intent to *monopolize* as the criterion of guilt. The minority judges contended that it was not the intent of Congress to declare *every* contract or combination in restraint of trade illegal, but only such contracts and combinations as aimed at monopoly — unreasonable contracts and combinations in restraint of trade.

This division in the court continued to appear in several decisions extending over a decade. Finally, in 1911, in the American Tobacco[1] and Standard Oil[2] cases the personnel of the court had changed sufficiently so that the former minority construction now became the ruling of the court. These two combinations were dissolved not on the ground that *every* contract or combination was illegal under the Sherman Act, but on the ground that evidence showed that they were consciously and deliberately *aiming at monopoly.* The combinations therefore constituted an unreasonable restraint of trade. Moreover, the court held that the certain proof of intent to monopolize was to be found in the long list of unfair practices by means of which these combinations had sought improperly and illegally to kill off competition. Such unfair practices are indubitable proof of intent to monopolize.

---

[1] 158 U.S. 564. [2] 166 U. S. 290. [3] 171 U. S. 505.

Thus the so-called rule of reason came finally to be applied to the Sherman Law. The law was interpreted to apply only to those contracts and combinations which are unreasonable; that is, which aim to monopolize. The use of unfair practices was held to be a clear proof of intent to monopolize.

Efficiency versus Unfair Practices. Let us suppose that a combination did not use unfair practices, but was able to attain a dominant place in the industry through efficient methods, that the competitors were finally pushed to the wall not by unfair practices, but through the superior efficiency of the combination. Such a combination might attain a monopoly through its superior efficiency in competition without aiming or intending to monopolize the field. Would such a combination be guilty of unreasonable restraint of trade?

In 1920 the United States Supreme Court gave a partial answer to this question in the decision of the United States Steel Corporation[3] case. Here was a holding company which at the inception of the combination in 1901 controlled about 60 per cent of the steel industry. As time went on, however, its control decreased to little more than 50 per cent. Here was a combination which controlled at least a half of this entire industry, but which was not using unfair methods against its competitors and was not, according to the findings of the court, pursuing a price policy contrary to public welfare. The court therefore refused to sustain the plea for dissolution. The practical effect of this decision is to permit combination, even though the combination assumes huge proportions, provided, first, that unfair methods are not used against competitors, and, second, that no public interest would be served by dissolution. It would seem, therefore, that we are tending in the direction of the German law with reference to combination. In Germany the Kartell agreements are valid contracts enforceable at law. But this does not mean that a Kartell may exploit the public by means of an extortionate price policy without being subject to prosecution in the courts. It is all a question of reasonableness. And it is this test of reasonableness that our Supreme Court has applied in the Tobacco, Standard Oil, and Steel Company cases.

---

[1] 221 U. S. 106.

[2] 221 U. S. 1.

[3] 251 U. S. 417.

It should not be inferred, however, that our courts are willing to go as far as the German courts in permitting reasonable combinations. Our courts are still skeptical of combination and strongly inclined to maintain competition wherever it is possible to do so. Pooling and trust agreements are still illegal in this country. Even certain "open price" associations have not been able to satisfy the courts as to the reasonableness of their activities. If, however, combination has been reached apparently as a result of natural economic tendencies without the use of unfair methods of attack upon competitors; if the combination appears to be economically efficient, so that in all probability the cost of production and the normal price would be higher were it broken up into less efficient competing units,— then the courts will take the position that the test of reasonableness should be applied, and the combination should not be dissolved as a matter of course without regard to questions of efficiency and expediency. It is here that the rule of reason comes in to modify the former arbitrary and mechanical construction of the Sherman Law.

Trade Associations. Recent decisions of the Supreme Court (rendered in 1925) in the Cement and Maple Flooring cases indicate a further tendency to interpret the Sherman Act liberally. As a result of these decisions it is lawful for trade associations to disseminate information with respect to prices that have been charged in different localities and to distribute statistics of production, stocks on hand, etc.

It must not be inferred from this, however, that a so-called gentlemen's agreement with respect to prices will be tolerated. As a matter of fact, in the earlier Hardwood[1] and Linseed[2] cases the court refused to recognize the legality of these trade associations partly on the ground that evidence existed from which it appeared that something like a gentlemen's agreement as to future production and prices prevailed in these industries. A trade association may, however, lawfully give its members information, but it is not permitted to engage in practices that will tend to establish production quotas or uniform prices.

The Federal Trade Commission and the Clayton Act. In this country there are many well-informed persons who hold that if unfair practices are prohibited, only highly efficient and economically sound combinations will be able to survive. It was on this theory that the Federal Trade Commission was established in 1914. Unfair practices are prohibited under the law, and the commission is intrusted with the duty of exposing and preventing them. It may make investigations on its own initiative or upon the request of the president, Congress, or the Attorney-General. Or in case of dissolution proceedings the courts may turn to the commission for a plan which will restore competition and fair practices and safeguard the interests of the public. If a company is found to be guilty of unfair practices, the commission may issue a restraining order which will be enforced if necessary through the circuit courts. Moreover, the Clayton Act sought to clarify the law as to what constituted unfair practices. Price discriminations, exclusive contracts, and holding companies are definitely prohibited where the effect is substantially to interfere with competition. Interlocking directorates are prohibited (1) between two competing industrial corporations, any one of which has capital, surplus, and undivided profits of a million dollars; (2) between two banks, any one of which is organized under the laws of the United States and has deposits, capital, surplus, and undivided profits of five million dollars[1]; (3) between railroads and railroad-supply companies, from whom railroads make purchases, unless such purchases are made under competitive bids supervised by the Interstate Commerce Commission.

---

[1] 257 U. S. 377.

[2] 262 U. S. 371.

This theory that inefficient combinations cannot withstand the pressure of potential competition if fair competitive methods are enforced finds considerable support in the history of American combinations. Professor A. S Dewing, who has made a study of thirty-five combinations organized in the main from 1898 to 1902, shows that twenty-two of the thirty-five earned less on the average during the ten-year period after the combination than had been earned by the separate plants before the period of combination. If we consider the entire thirty-five combinations it was found that the earnings before the combination period were between a fifth and a sixth greater than the average for the ten years after the combination. Moreover, the earnings before the combination were nearly a fifth greater than the earnings of the first year after the combination.

Illustrations of Monopoly based on Unfair Practices.[2] The American Sugar Refining Company is an example of a trust which in the end proved unable to withstand the continuous pressure of new competition constantly springing up on all sides. This corporation had the advantage of operating behind a tariff wall which practically excluded foreign competition from the early eighties until 1913. In the early nineties this company had succeeded in forcing (under the pressure of severe competition) a sufficient number of its competitors to sell out, so that it controlled 90 per cent of the manufacture of sugar in the United States. When the company had acquired a substantial monopoly control, the price was raised. Large profits attracted competitors. New refineries were built. The trust struggled to beat them down or else get control over them. Price wars were instituted to bring them into line. By 1900 most of them were brought into the combination. Prices were again raised, and new competitors again sprang up. Price-cutting was again resorted to, and competitors, face to face with bankruptcy, yielded to the trust. Had it not been for the growing importance of beet sugar it is possible that the trust might have been able to choke off potential competitors indefinitely by these methods. But it has been estimated that in 1919 less than 30 per cent of the total production of sugar was controlled by the American Company; however, if we include the independent companies in which the American Company has an important interest, the percentage would be considerably greater. Other unfair practices of which the sugar trust was guilty were (1) receiving illegal railroad favors and (2) defrauding the government in the payment of tariff duties on raw sugar by false weighing.

---

[1] This portion of the act was subsequently amended so as to permit interlocking directorates with not more than two other banks subject to the approval of the Federal Reserve Board, if in their judgment the public interest would not be injured thereby.

[2] See Eliot Jones's *Trust Problem in the United States.* The Macmillan Company.

The Standard Oil Company and the American Tobacco Company are illustrations of trusts that built up their immense power through the use of unfair methods of competition. Railroad favoritism and destructive price-cutting were large factors aiding the Standard to get a monopoly control of pipe lines. Once having acquired a monopoly, the Standard exerted its influence to prevent new pipe lines from getting the right of eminent domain and from crossing railroad tracks. Such new pipe lines as were built were immediately bought up, or the wells and refineries using these pipe lines were purchased and the business diverted to the Standard lines.

Sometimes, without going to the extent of purchase, the Standard brought pressure to bear upon the owners of these wells and refineries to divert their business to the Standard lines. Because of the monopoly control over pipe lines, independent refineries were charged excessive rates and their profits destroyed.

In the transportation of refined oil the Standard, from its early history, was able to secure railroad rates which were more favorable than those offered competitors. After rate discrimination was made illegal, secret rates favoring the Standard were obtained. Moreover, the published rates favored localities in which Standard refineries were situated.

Another unfair method by which monopoly power was achieved was the practice of reducing below cost the price of refined oil in localities where competition was intense. The purpose of these price wars was to drive the competitors out of business and later recoup losses sustained in the fight by establishing monopoly prices. A great national concern like the Standard was well equipped to engage in these local price wars, since it could balance losses in competitive areas by monopoly profits elsewhere. In this struggle against competitors the Standard was aided by a well-developed spy system.

The American Tobacco Company is another classic example of a trust built up by unfair methods. The company was organized in 1890 after a period of severe competition between leading cigarette manufacturers. Exclusive control was obtained over patents and machines used in making cigarettes. Competitors were bought out at excessive prices, and the plants thus acquired were frequently closed. By these means the trust acquired a control over the cigarette output amounting to 93 per cent in 1900 and to 86 per cent in 1910.

A few years after it was organized, the company decided to get control of the plug-tobacco business. Ruinous price-cutting forced most of the independent plug manufacturers into line. Other companies were purchased at excessive prices. By 1900 the trust controlled over 60 per cent of the business, and this figure was increased to 85 per cent by 1910. Similar tactics resulted in an 80 per cent control of the snuff business by 1901, and this control increased to 97 per cent by 1910. The trust also secured a 75 per cent to 95 per cent control of smoking tobacco, fine-cut tobacco, and licorice. Only in cigar-manufacturing was it unable to secure a dominant position. Machinery is relatively unimportant in the manufacture of cigars. Hence an enormous number of small concerns were able to flourish. Without large-scale production and concentration of ownership, combination is difficult if not impossible.

The Tobacco and Standard companies were dissolved by court decree in 1911 because of their monopoly position and the unfair methods by which they had secured and continued to retain their power. The Federal Trade Commission in a subsequent investigation (1915) found that the dissolution decrees were not appreciably effective in restoring competition. In each case the trust was broken up into a dozen or so separate companies, but community of interest still prevailed because of common ownership of the stock in the different companies by a small group of men. The several Standard companies divided up the territory, and as a result there was no competition in the marketing of gasoline. The field was divided into eleven distinct marketing areas which followed state lines and were therefore obviously arbitrary divisions. The commission found abnormal inequalities in the price of gasoline in the different areas. It also found that the Standard companies controlled about 65 per cent of the gasoline business, and this percentage would be increased if certain so-called independents (controlled in part by the Standard interests) were included. The commission found that on the whole the independents followed the prices fixed by the Standard companies. The commission recommended legislation which would make it possible to reopen trust cases where conditions apparently justified a modification of the dissolution decree. In this case it would be necessary to eliminate the community of interest between common stockholders and thus separate the corporations into real competing units.

The dissolution of the tobacco trust apparently increased the competition for the business among the several companies; for there was a considerable increase in advertising expenditures. The dissolution, however, did not result in lower prices.

Quasi-Monopoly based on Efficiency. On the other hand it appears that on the whole the International Harvester Company and the United States Steel Corporation have not built up their monopoly or quasi-monopoly power through unfair methods of competition. Having already given some of the leading facts with regard to the Steel Corporation, we shall confine our discussion here to the Harvester Company.

After a period of strenuous but not ruinous competition, five leading manufacturers of harvesting machines combined and organized the International in 1902. They controlled 85 per cent of the output of harvesting machines, 83 per cent of the output of mowers, and 80 per cent of the output of rakes. The control was in reality greater than this, for the independent companies were confined mainly to the Eastern market. By 1918, however, the control had dropped from 80–85 per cent to 60–65 per cent. Other lines were also entered. By 1910 the company had a 50–55 per cent control of the output of harrows, manure-spreaders, corn-shredders, and binder twine. Other lines in which the trust entered, but in which no monopoly control was acquired, included wagons, plows, trucks, seeding machines, gasoline engines, tractors, cream-separators, cultivators, hay-loaders, and hay-stackers. The company also acquired timberlands, coal lands, iron-ore properties, iron-and-steel mills, and railroads.

Though the Harvester Company had not used unfair methods such as were practiced by the oil and tobacco trusts, yet because of its monopoly position the government brought suit in 1912 with a view to dissolution. The Federal Circuit Court handed down a decision in 1914 calling for dissolution on the ground that 80–85 per cent control constituted an unreasonable restraint of trade. Judge Sanborn, however, dissented. He held that unreasonable restraint involved (1) raising prices, (2) limiting production, (3) deteriorating quality, (4) decreasing the prices of labor and materials, (5) oppressing competitors. He did not find the Harvester Company guilty under any of these counts and therefore held that the company did not unreasonably restrain trade. The case was appealed; but in 1918, before a decision was given, the company agreed to a consent decree, calling for the sale of the Osborne, Champion, and Milwaukee plants. The Federal Trade Commission expressed the opinion, however, that an effective dissolution would involve more drastic action than was required in the consent decree, and that what was needed was the breaking up of the McCormick, Deering, and Wisconsin Steel companies into separate and distinct companies with separate stockholders.

The Harvester and Steel companies thus appear to be able to hold their own because of efficiency without having to resort to unfair methods of attack on competitors.

Monopoly based on Patent Rights. The United Shoe Machinery Corporation is a trust whose monopoly power rests neither on unfair practices nor on efficiency, but on patent rights. The corporation is a combination of several companies, each of which had patents on shoe machinery covering different processes in the manufacture of shoes. The government instituted a suit against the corporation, claiming that it constituted a combination in restraint of trade. The case was decided by the Supreme Court in 1918. The decision hinged on whether the patents owned by the several companies were complementary or competitive. The court decided that the trust was not illegal since each of the various companies entering into the combination constituted a patent monopoly. The court held that the patents covered different processes, were noncompetitive, and that therefore the monopoly power of the combination was no greater than that of the separate companies. The dissenting judges were of the opinion, however, that the several companies were in part competitive; that the contracts under which the corporation leased its machines constituted a restraint of trade in that they restricted the lessee exclusively to the use of the corporation's machines. The majority judges, however, held that such tying clauses go with patent rights.

International Kartells.[1] Agreements between the producers of different nations had already reached considerable proportions in the nineties, and the movement continued to increase up to the time of the World War. Many agreements were disrupted as a result of the war, but in recent years the movement has revived on a large scale. These agreements commonly assign the home market to each national group and frequently make sales allotments or establish prices in third countries. The agreements may be made between various national Kartells or combinations, or they may be made directly between the individual firms in different countries, or they may be made between national Kartells or combinations in some countries and individual firms in others.

Numerous international Kartells have existed from time to time and still exist in the chemical industry. Other industries in which important Kartells have been organized are steel rails, enameled ware, cement, aluminum, watches, and ocean shipping.

Many international Kartells have been organized or reorganized since the war. In 1925 an international agreement was made in the incandescent-lamp industry. The General Electric Company is an important factor in this industry. In Europe it operates through a subsidiary, the International General Electric Company, with head offices in Paris. The Kartell includes syndicates and firms operating in France, Germany, England, Italy, Spain, the Scandinavian countries, and Holland.

In the copper industry there is an international syndicate which is said to control 90 per cent of the world's production of copper. Included in its membership are the American Smelting and Refining Company, the Kennecott Copper Company, and the Anaconda Company. These three American companies alone control over 50 per cent of the world's output. Producers in Germany, South America, and Africa are also included. The central office of the organization is the Copper Export Trading Company in Brussels. Several sales agencies are maintained in England, Germany, and France. Among the independents not included in this world syndicate are the Miami Copper Company (United States), a Spanish-British company, and the Australian producers.

---

[1] See Robert Liefmann's article "International Cartels," in the *Harvard Business Review*, January, 1927.

In the crude-steel industry Kartells have been organized including producers in Germany, France, Belgium, Luxemburg, Austria, Hungary, Bohemia, Poland, and Rumania. In 1925 a German-French Kartell was organized in the potash industry, and more recently one was organized in the steel industry.

International Concerns. In a few industries certain firms have acquired a dominating position throughout the world. The most striking illustration of such international concerns, and one that comes very near to being a world monopoly, is the Swedish Match Company, which controls plants in thirty-five different countries either directly or through its subsidiaries. It owns the controlling stock in the International Match Corporation, an American company, which in turn controls companies owning ninety plants in various European and other countries. The Swedish company also controls the German Match Manufacturing Syndicate. Moreover, through the advancement of loans, it has secured the coöperation of a large number of state match-manufacturing monopolies, including France, Poland, Portugal, Greece, and Peru.

Similarly, in the petroleum field the Standard Oil Company of New Jersey has reached an outstanding international position. It is more or less associated with numerous American companies, generally known as the Standard Oil group; it is interested in some two hundred and fifty companies outside the United States; in Europe it has made numerous agreements with Kartells with respect to territory, prices, and sales quotas.

In the field of electrical products the General Electric Company is rising to increasing importance. In 1920 it acquired a large block of shares of the German General Electric Company. These two companies, moreover, made agreements in 1924 with respect to exchange of patents.

Price Policies. The trade practices of the various industries with respect to price policy throw light on the way competition works in the modern organized economy. In many industries the usual practice is to sell f.o.b. at the mill, the purchaser paying the cost of transportation from the mill. This practice comes nearest to satisfying the requirements of a truly competitive market. Under this system the buyers are fully enabled to exercise alternative choices and to purchase at the lowest-price plants, cost of transportation considered. This policy, universally pursued, would result in the most economical geographical distribution of industries and hence in the most efficient utilization of the nation's resources.

f. o. b. mill

In certain industries it is the practice to sell the products f.o.b. point of delivery. This would tend to establish a uniform price at all points regardless of the distance from the producing plants. The Portland cement industry follows this practice. The practice is facilitated somewhat by the fact that cement is produced in about one hundred and thirty plants scattered in various parts of the country. Many trade-marked articles are sold at a uniform price in all parts of the country.[1] Other industries pursue the policy of dividing the country into zones and quoting a special f.o.b. price at point of delivery in each zone. This was at one time the practice of the steel companies. The zones radiated from Pittsburgh as the center. The more distant the zone the higher the delivered price.

f. o. b.
Point o
deliver

Zones

This uniform-delivered-price practice is discriminating and wasteful. If the price in any given zone or territory with a uniform-delivered-price system is a fair average price, it is clear that consumers located near points of production are being charged unduly high prices in order to make it possible to sell to distant points below cost, including cost of transportation. If the commodity can be produced at the distant points as cheaply as at the points of production, then plants ought to be located there and thus save the cost of transportation; if it cannot be produced as cheaply, the uniform-price system is wasteful from a national standpoint, since it places regions on an equal footing which in fact are not equal. The effect is to stimulate artificially regions that are handicapped as to location or natural resources. Such artificial stimulation results in an uneconomical utilization of a nation's resources.[1]

---

[1] Similarly, department stores deliver to all parts of a city at a uniform price.

Some industries quote a delivered price based on the price at a certain designated point plus cost of transportation to the point of delivery. Until 1924, when the Federal Trade Commission ordered the practice discontinued, the steel industry followed this practice. This was the "Pittsburgh-plus" system, Pittsburgh being the basing point. This practice clearly resulted in discrimination against buyers who purchased from plants located outside of Pittsburgh. Thus steel consumers in Chicago had to pay the Pittsburgh price plus cost of transportation to Chicago, even though they purchased the steel from a Chicago mill. The Chicago steel mills sold steel to consumers in Ohio at the Pittsburgh price plus transportation from Pittsburgh to Ohio. This price would be much lower than the Chicago price. Clearly the Chicago consumers were discriminated against in this case, especially in view of the fact that the Chicago mills were said to be the lowest-cost mills. A freely competitive market ought to permit consumers to buy at the several mills, thus enabling them to compare one mill price with another and to buy in the most favorable market, transportation costs considered.

---

[1] In some industries (department stores, for example) it is simpler and probably more economical to charge delivery to overhead instead of making a special delivery price for each commodity.

Many industries follow the practice of quoting different prices in different territories in order to meet competition. This is not considered in law to be an unfair practice if it is done to *meet* competition and not to *destroy* it. In certain cases a discriminatory price is economically justifiable. In the railroad business it is often advantageous to charge less for a long haul than for a short one. If the long-haul traffic cannot be obtained at other than exceptionally low rates, and if such traffic nets something toward the fixed expenses over and above the mere cost of hauling, then the low rates can be justified. Were it not for this additional traffic the short-haul traffic would have to pay still higher rates to cover the total expenses of running the railroad system. But this is true only in so far as the railroad is not being used to full capacity by normal-rate traffic. As the normal-rate traffic grows it will, of course, be economical to discard the low-rate traffic.

In the case of a manufacturing industry a firm can economically dump surplus production at distant points or in foreign 6
countries, provided the surplus can be produced at a cost which is lower than the price obtained; in short, "dumping" is economical so long as the plant or firm has not yet reached the most efficient size, so long as *additional* output incurs only a slight *added* cost. Suppose, for example, a plant located in Minneapolis finds that the market naturally tributary to it is not large enough to permit the most economical large-scale production. The Minneapolis plant will then naturally seek to capture more distant markets by laying down a lower price than its competitors, perhaps lower even than the Minneapolis price. Price discrimination of this sort is justifiable in so far as such a policy results in larger-scale production at a cost sufficiently low to offset the higher cost of transportation that is necessarily incurred as the market is widened.

The Modern Organized Economy. Wherever artificial-price policies are in effect, free and untrammeled competition does not prevail. Perfect competition and perfect monopoly are alike rare in the business world today. Between the two extremes there are all sorts of gradations. Few products are under strict monopoly control, but, on the other hand, few products reach the consumer without being subjected to more or less restraint of competition.

Innumerable goods are protected from the leveling effects
1 of perfect competition by brands and trade-marks as well as
by patents and copyrights. In many cases the manufacturer
2 fixes the prices to be charged by retailers, thus eliminating
competition between retailers as to price. Trade associations
build up a sort of "trade ethics" which tends to prevent
3 underselling. Open-price associations serve to standardize
prices by supplying their members with information as to
costs and prevailing selling prices. Trade unions affect the
4 cost of production and hence the prices of the products in
fields in which they are strong. Farmers' coöperative market-
5 ing associations attempt to adjust supply to demand and at
times probably have an influence on total production.

Let us take the milk-producers' associations as an illustration of the effect of organization on competition and prices. The Twin Cities Milk Producers' Association is an organization composed of farmers who furnish milk for the Minneapolis-St. Paul market. It controls 75 per cent of all the milk produced in this area. Its substantial monopoly of the milk supply gives it a strong position in its bargaining relation with the distributors' association; yet this position by no means frees it from the pressure of competition as a factor in fixing the price. To begin with, it is not impossible for the distributors to ship large quantities of milk in from considerable distances. Moreover, it would not be to the best interest of the milk-producers themselves to raise the price of milk unduly, for the reason that an excessive price would reduce the quantity sold. The problem would then arise as to what to do with the surplus, because the association does not aim to limit production. The surplus could, it is true, be used to produce cheese or butter, which could be shipped into the general markets. If milk-producers' associations did this sort of thing all over the country, however, the effect would be to depress unduly the prices of cheese and butter; thus what would be gained on milk would be lost on butter and cheese. Furthermore, such a situation would tend to divert competitors from the production of cheese and butter to the production of milk. It is true that if such a policy were followed in only two or three cities, the milk-producers might profit considerably by charging an abnormally high price for milk and converting the surplus into butter and cheese for the general market. But two reactions would inevitably follow: (1) distant producers would ship in their milk, and (2) an aroused public would force the government to proceed against the association as a combination in restraint of trade and contrary to public policy.

It is a great mistake, therefore, to assume that the existence of organized groups in our modern economic life means that the pressure of competition is no longer a pervading one.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The dairy farmers within a forty-mile radius of the Twin Cities are organized in the Twin Cities Milk Producers' Association, and the distributors are organized in the distributors' association. Thus there is a double monopoly, leaving the unorganized consumers entirely helpless." Do you agree? What factors safeguard the interests of the consumers?

2. "Since farmers' marketing associations are protected by law, there is no legal recourse against monopoly prices of products sold by them." Do you agree?

3. "Farmers' marketing associations are not monopolies, because they do not restrict output." Give arguments for or against this view.

4. "All kinds of private businesses are being bought up by banking houses. . . . The net effect . . . is the assumption of an absolute control of intermediaries — most commonly bankers, so-called — in place of the former responsibility for direction which, theoretically at least, rested upon the shoulders of the actual owners." — RIPLEY

What, if any, is the bearing of this "insinuation of banking power between ownership and operation" on the monopoly problem?

5. "The recent Dodge Brothers, Inc., is typical. A banking house buys up a private business for, let us say, $146,000,000. This sum and more, it recovers by the sale to the public of $160,000,000 of bonds, preferred stock, and 1,500,000 nonvoting shares of 'Class A' common stock. But not a single one of the 500,000 'Class B' voting common shares is thus sold. The promoters have virtually paid themselves a handsome profit for the assumption of the entire directorial power, having mortgaged the property to the full amount of its original cost, including both assets and capitalized earning power." — RIPLEY

Is there any reason to suppose that such manipulations will injure the consumers? Why, or why not? Are such developments likely to lead eventually to monopoly or quasi-monopoly in the automobile business? Give reasons for your answer.

6. "Henry Ford has reached a position where he wields tremendous monopoly power." Argue for or against this statement.

7. In 1920 a consent decree was entered into between the government and the five leading packing companies. The packers had obtained not only a dominant position in the packing industry proper, but also owned the leading stockyards and had entered extensively into the wholesale-grocery business. Under the terms of the consent decree the packers agreed to withdraw from the wholesale-grocery business and to sell their interest in the stockyards. Later the California Coöperative Canneries Company asked to have the decree modified so that the packers might continue to distribute their products. This was objected to by the Southern Wholesale Grocers' Association and the National Wholesale Grocers' Association.

Would the entrance of the packers into the grocery business mean more competition, or less, from the short-run point of view? in the long run?

8. Is it likely that the recent combination of baking companies will be able to establish a bread monopoly? Why, or why not?

CHAPTER XVIII

PUBLIC REGULATION OF MONOPOLY PRICES; PUBLIC UTILITIES

The Nature of a Public Utility. In the preceding chapter the efforts of the public to prevent the formation and continuation of monopolies among various types of business have been discussed. Generally the attitude of lawmakers and of the public has been that monopoly is harmful to the consumers and to those who wish to compete in the business which becomes monopolized. In certain types of production, however, monopoly is either directly fostered by public authority or is recognized as being the inevitable result of the economic conditions under which production is carried on. Aside from monopolies based on patents and copyrights, the principal examples of recognized monopoly are found in the field of public utilities. Here the chief representative at the present time is the railroad. But electric power, street transportation, the supply of gas and water, the telephone and the telegraph, and sewage disposal are also included. Sewage disposal is universally a free service owned and controlled by municipalities. Water supply is also under public ownership in a majority of our cities. These services are so closely involved with the necessity for safeguarding public health that the public usually does not permit them to be privately owned. There are numerous exceptions to this rule in the case of water supply, however. All the other public utilities can be privately owned, and in the United States they are predominantly in the hands of corporations whose objective is business profit.

The legal definition of a public utility is easily stated but difficult to apply in particular cases. In general any business becomes a public utility when it can be shown that it is "affected with a public interest." Reduced to simplest terms, a business is affected with a public interest if it supplies a very important commodity or service (but it need not be a necessity of life) and when there is danger of monopoly of the service. When these conditions exist the legislature may step in and impose regulations governing the kind and quality of service and the prices that may be charged for it. The immediate decision, whether importance of the service and danger of monopoly are present to a degree that warrants regulation, is the function of the legislature; but in the United States the final decision rests with the Supreme Court of the United States. Historical precedent plays a part in these decisions. Railroads have always been held to be affected with a public interest, and this has been true of all municipal utilities that require special permits to use public streets.

If we go back several hundred years in the history of English institutions, we find that the charges of a large number of businesses were subject to regulation. One reason for this condition was the greater likelihood of monopolization. The absence of cheap transportation and the lack of communication, the lack of adequate methods of finance provided through the modern banking system, and the greater prevalence of legal and customary restrictions on the process of entering a trade made the establishment of local monopolies much easier than in the early years of the present capitalistic era. With the widening of the market and the growth of the competitive spirit in business, both of which accompanied the Industrial Revolution, the amount of governmental regulation declined. Freedom of enterprise came to be regarded as essential to efficient production and to economic progress.

But with the development of capitalism the nature of competition gradually became different from what it was in the early nineteenth century. The use of large amounts of specialized capital made competition more dangerous to the competitors; the economies of large-scale production limited the number of competitors in many industries and made combination easier; and the development of modern financial institutions, such as the corporation, the stock exchange, and the banking system, facilitated the acquisition and retention of control of competing concerns. These forces have tended to restrict competition in many kinds of production, and they have affected public utilities as well as manufacturing and merchandising. Railroads have combined into large systems, one company controls the major part of the telephone service of the entire country, and the supply of gas and electric current is dominated by a small number of corporations, which are connected through holding companies and other intercorporate relations. Yet, although the organization of holding companies and the development of other methods of combination in the field of public utilities are undoubtedly but a part of the general movement toward centralization, there are certain economic reasons why a public utility tends to become a monopoly within the territory which it serves.

Economic Characteristics of Public Utilities. In the first place, public utilities are likely to become monopolies because the principle of increasing returns operates strongly in the production of such services. As has been emphasized in the discussion of decreasing cost, a great density of traffic is required to enable a railroad to bring its unit costs to the lowest point. The same is true of the supply of electric current. In both these utilities the amount of fixed, specialized capital is very great, and the possibilities of economy from the use of large mechanical units are numerous. On the other hand, a telephone exchange soon reaches its most economical size, and thereafter costs increase as more and more subscribers are connected with the system. As far as the manufacture of gas is concerned, decreasing cost seems to play a minor rôle.

The tendency toward monopolization of public-utility services is also due to peculiar limits on their power to compete for business. A railroad, a street-car line, a gas company, or a telephone system can only offer service in the restricted territory where its lines are located. A manufacturer is, by contrast, not so limited. If he is undersold in one territory, he may find other markets for his wares without much additional expense. And the merchant who finds a town oversupplied with stores can transfer to another town. The inability of utilities to compete is also restricted by the absence (6) of variety in the commodities or services rendered. The manufacturer whose products are being driven from the market by the price-cutting of rivals can usually change the type of good he turns out and appeal to different demands or to different groups of consumers. The same thing is true of the merchant. But a railroad must carry freight and passengers or quit business entirely.

A third peculiarity of public utilities is the fact that to 3 render the services required of them they must encroach upon the rights of other users of land. A street-car line interferes with automobile traffic and it depreciates the residential value of adjacent land. The supplies of gas and water require subsurface piping, which is usually laid in the public street. The laying and repair of mains disturbs the surface of the street and interferes with traffic. Telephone and electric-power systems make use of wires which interfere with fire protection and detract from the utility of residential property. Because of these encroachments on the utilization of land by other property-owners and by the public, municipal utilities would probably be given exclusive rights within certain areas even if there were no other forces making for monopolization. The duplication of utility systems is inconvenient to the public and is also useless and wasteful.

No single characteristic of public utilities is sufficient to account for monopolization under all conditions. In areas of dense traffic railroads are competitive businesses, and they might remain so even if the public should withdraw all regulation and allow them to make their own arrangements; or at least there would be no more tendency for them to combine than there is for manufacturers, whose plants require a large amount of specialized capital, to enter into agreements to restrict competition. Urban utilities which carry people about the city do compete at the present time, as is well known. In large places suburban railroads, elevated roads, subways, and surface lines are all more or less in competition. In all large cities the bus lines compete with the street cars, and they compete with the railroad for intercity business. Different concerns often compete for the supply of electric power to manufactures, although there is little tendency for competition for lighting to persist. Monopolization is usually found where the demand for a given service is not sufficient to enable two competitors to produce at the lowest average unit cost, or where public convenience forbids duplication of distribution systems (or service lines in the case of street cars), or where a combination of these conditions is present.

The public regulation of public utilities has two principal objects in view: (1) the regulation of charges to prevent the exploitation of the consumers and to prevent discrimination between different consumers; (2) the regulation of services regarded as vital to the health, the safety, or the general welfare of the community. The latter is chiefly a technical matter; and though it has many economic bearings, we shall confine further discussion of regulation to problems of public-utility rate-making.

Purpo
of

Regulation of Public-Utility Earnings. The making of public-utility rates is an exceedingly complex problem. Here we shall not attempt more than a general treatment of the subject, and in this treatment we shall go only as far as is necessary to show in a typical case — that of railroad rates — the operation of economic principles in public regulation of this type of monopoly.

The first problem which confronts public authorities is the amount of income that the property of public-utility corporations shall be allowed to gain from the services they render. In court decisions it is commonly held that the investors in public utilities are entitled to a "reasonable return" on the "fair value" of the property devoted to the service of the public. This could easily be applied if the two factors "reasonable return" and "fair value" could be accurately defined. The determination of reasonable return involves relatively little difficult economic reasoning, although the actual application of the definition is often more complicated. In general it is agreed that the investors must be allowed a return which will equal the return they could obtain in competitive businesses of equal risk. The reasoning by which this principle is established is simple. If the return is below the level of the return in other businesses, the utility corporations will be unable to obtain new capital for expansion, and the public will suffer from lack of railroad service, street-car service, and so on. If the return is above the competitive level, the consumers are being exploited by the monopoly, and the investors are receiving more than is necessary to induce them to offer capital for expansion. It has been decided by the courts in many cases that to fail to allow the investors such a fair return on their investment is a taking of property without due process of law and therefore in violation of the Federal Constitution. In this instance the legal reasoning and the economic reasoning are in agreement.

sonable

The principal difficulty in the determination of the earnings to be allowed a public utility is the finding of the fair value of the property devoted to the service of the public. As yet neither the courts nor the economists nor the experts in the management of the utilities have been able to agree upon a practicable rule for the determination of fair value. Within each of these groups there is still further lack of agreement. Here we are concerned primarily with the economic aspects of the different attempts to set up a rule for the guidance of public officials who are required by law to establish the rates which will allow only a fair return on the fair value of public-utility properties.

Market Value of Utility Property. It was once suggested that the same rule should be applied to these properties that is frequently made use of in the valuation of competitive businesses.

The value of the agents of production depends on their prospective earnings and on the rate of interest; in other words, the sale value of a factory is its present worth as a going concern. If all parts of the factory are reproducible, this value will not, of course, be very much above the cost of reproduction. But it may be very much below that cost if the plant contains some out-of-date equipment, or if the industry is overbuilt, or if the plant is improperly located, or if the management in the past has been inefficient. But, whatever the cost, the present worth of prospective earnings is the fundamental factor in the determination of the market value of competitive businesses.

This rule is, however, not applicable to public-utility properties because they are monopolies. In the absence of public regulation their earnings are not limited to a fair return on the competitive value of the property. But these earnings will be capitalized at the rate of interest current in competitive business. The result is that without regulation utility properties may have a value as going concerns far above the competitive value of the agents of production they employ. But one purpose of regulation is to prevent the earning of monopoly revenue by the utility; and this could not be done by allowing a corporation that is being brought under regulation for the first time to capitalize its monopoly profit.

When regulation has been in force for some time, the earnings of the corporation depend upon the rates which the regulatory body has imposed. To capitalize these earnings for the purpose of ascertaining the fair value of the property is to assume that regulation in the past has been correct. But the question before the regulatory body is to determine whether the regulation has or has not been correct. To take the market value of the property or the market value of the stocks and bonds of the corporation as the fair value upon which the corporation shall be allowed a fair return is to reason in a circle.

Original Cost. Another theory holds that the fair value of utility property can only be established by getting at the

(c) original cost of the property minus depreciation due to obsolescence and to wear and tear. How much money was actually invested in the right of way, in roadbed, ties, rails, engines, cars, stations, etc.? Regardless of the securities that were issued, how much money was actually paid over to the corporation and invested in railroad property? From the standpoint of this theory fair value is the actual investment value minus depreciation.

At first thought this appears to give an easy and satisfactory solution, but further examination shows that the problem is not so simple. To begin with, certain practical difficulties at once present themselves. The actual investment can seldom be accurately determined, simply because accounting records are not complete. A great deal of data of this sort has been gathered by the Interstate Commerce Commission, but it is doubtful if it is accurate enough to give a complete valuation of railroad property from the standpoint of original cost.

Moreover, certain theoretical difficulties appear. Are we to include, in the fair value of the railroads and other utilities, property derived from excessive earnings? Clearly an actual investment was made in this property, but it came out of the earnings of the corporation and not out of individual investments. If the excess earnings were due to exorbitant rates, it has been argued that we should not permit the reinvested capital so derived to be counted as part of the fair value of the property. According to this point of view, if we permit the utilities to make a fair return on capital derived from past unreasonably high rates, we are reasoning in a circle just as truly as though we attempt to arrive at a fair value on a market-value basis. If we permit a high valuation as a result of past high rates, then we are allowing the rates in the future to remain high because the rates in the past were too high.

But now let us suppose that these excessive earnings have been distributed to the stockholders. This, of course, would be entirely legal. The stockholders, let us suppose, buy additional securities with the funds so distributed, which the utility in turn invests in additional plant. Shall this property be allowed as part of the fair value? It is difficult to see how it could reasonably be excluded. At any rate, as a practical proposition, it would be absolutely impossible to do so, since no one could discover the source of the individual incomes invested in these securities. But if these excessive earnings are to be counted as part of the fair value when reinvested by stockholders, why not count them when reinvested by the corporation itself? There is, however, this difference: if, indeed, a high-dividend rate had been established, it is not improbable that the public would have insisted on a reduction in rates, thus preventing the excessive earnings. So long as the surplus earnings were being put back into the business the public was not fully aware of the fact that the earnings were really excessive. Nor is it sufficient defense to say that this surplus is the property of the utility corporations, for no property right is unlimited. The public, which has provided this surplus through the payment of excessive rates, has an interest in the property. This argument is greatly weakened, however, if the rates in question were already regulated by the government. The Interstate Commerce Commission has, in fact, passed upon this question in two important cases. The Chicago, Burlington and Quincy Railroad and the Delaware, Lackawanna and Western Railroad asked to be permitted to capitalize the reinvested surplus by an issue of stock dividends. The Commission granted these requests, declaring that there was no proof that the surplus had resulted from excessive rates, since the rates had been controlled by regulating bodies and were substantially the same as those charged by competing lines.

Even though there had been no rate regulation the Commission contended that the surplus could properly be capitalized. It said: "Where the public has found it expedient to adopt a laissez-faire policy to encourage utility development, it cannot be said that profits have been illegally collected in the absence of regulation. The title to the surplus has vested without limitation or condition in the corporation and benefits the shareholder. The doctrine of implied trust sometimes applied by courts and commissions to donated property has no application to excessive return, for the payment of rates carried with it no requirement that the funds be left in the business or used for the public benefit. Its strained application to carriers who have made additions and betterments from surplus would only penalize those who came nearest to benefiting the public. The surplus from income was unrestricted legal property of the company, and ceased to be funds of the public, before the decision to divert it to either dividends or additions and betterments was made."

Similarly, in the Consolidated Gas case in New York City the court held that profits made under the supervision of a regulating commission were presumably reasonable and the earnings therefore had become the private property of the corporation. On the other hand, prior losses cannot be included in the valuation entitled to a return.

This brings us to the difficult question of the bonuses of money and land which the railroads received from the various governments — local, state, and Federal. Most of the land has since been sold and a part of the proceeds has been invested in railroad property. Such property certainly constitutes an actual investment, but the investment was not made by the stockholders and bondholders but by the public. It has therefore been argued by those who support the original-cost theory that property so derived should not be included in a fair value of the property, for although these gifts were made to help the railroads in building their lines, it was never the intention of the public to allow the earning of income on property derived from these gifts. The lands were not granted to enrich stockholders, but to help toward building needed railroad systems. The stockholders should not be permitted, it is said, to enrich themselves out of revenues derived from railroad property purchased from the proceeds of the sale of these lands. Such property may be said to belong, in a sense, to the public, since it was obtained from public funds; and therefore the railroad, it is argued, should not be permitted to earn any revenue on this property. Such property should therefore not be counted as part of the fair value on which a fair return may be allowed, since the security holders made no actual investment in this property.

On the other hand, it may be said that these lands were granted to induce individuals to invest in railroad shares when the roads were to be built through new territories. They were of little value at the time, but it was expected that they would become valuable in the future. The stockholders were expected to take small returns, or losses even, during the early years and make them up when the lands became valuable.

Another difficult question connected with the original-cost theory has to do with fluctuations in the value of money. If we take the original cost as the fair value, shall we consider the *number* of dollars actually invested or shall we consider the *purchasing power* of the dollars invested? If we take the number of dollars, a rise in the price level will result in a large confiscation of railroad property, and the real return will be far below a fair return on the purchasing power of the investment made. In the Southwestern Bell case the United States Supreme Court held that the existing rates were confiscatory, in part on the ground that no allowance had been made for the great advance in prices. The position of the court with respect to this matter is, however, by no means clear; nor is there agreement among legal writers, economists, accountants, and engineers. The Interstate Commerce Commission has, however, in the St. Louis and O'Fallon case, settled upon the 1914 price level as the basis for its tentative valuations.

Cost of Reproduction. In the St. Louis and O'Fallon case the railroad advanced the cost-of-reproduction theory. According to this theory the fair value is to be determined not by what the property originally cost, but by what it would cost to replace it now. As now applied in rate cases this theory occupies a ground quite the opposite of the original-cost theory. It disposes of the price-level controversy by adopting present cost instead of the original cost. It obviously includes in the fair value the reinvested surplus as well as railroad property derived from the sale of the land grants. It does not look into the source of railroad property, but asks the question What would it cost to replace this property now?

This theory, however, encounters new difficulties. Shall we take the cost of replacing the existing property by new property throughout, or shall we allow only the cost of replacement minus depreciation? Many railroad officials contend that nothing should be deducted for depreciation, since renewals and upkeep constantly keep the road in as fit condition as though it were new. However, in 1907 the Interstate Commerce Commission made provision in the classification of operating expenses for depreciation on equipment, and in 1914 provision was made for depreciation of fixed property as well as equipment. Moreover, in its valuation reports it has taken account of depreciation of the various items of railroad property.

In the opposite direction, what shall be done with property that has *appreciated* in value? This question is of special importance in the case of land. According to the original-cost theory, appreciation of land value should not be taken account of, but the cost-of-reproduction theory includes appreciation in determining fair value. This would permit all utilities to charge higher rates merely because the lands they use have gone up in value. In its tentative valuations the Interstate Commerce Commission has included present value of common-carrier lands (present value being estimated from the present average market price of adjoining lands) as part of the fair value.

In the Valuation Act passed by Congress in 1913 the Interstate Commerce Commission was directed to consider the original cost to date, the cost of reproduction new, the cost of reproduction less depreciation, and all other elements of value, if any. The Commission organized a Bureau of Valuation, which proceeded to take a census of all railroad property. This enormous task has not as yet been completed. The Commission, as we have seen, has issued a number of tentative valuations which appear to adopt, in large part, the cost-of-reproduction-less-depreciation theory, and it has set a tentative valuation of all railroad property at $18,900,000,000.

Specific-Rate Principles. We now turn to the question of specific rates; that is, the rates on each type of service. We speak of the general level of rates just as we speak of the general level of wages. Similarly, specific rates have to do with the rates on each separate shipment in the case of railroads and on each class of service in the case of other utilities, just as specific wages have to do with the wages of each type of labor in the different localities. From an economic standpoint the general level of rates must be high enough to give a fair return on a fair value, or, in other words, the total revenues must be sufficient to pay all costs, including a fair return on invested capital. It might be supposed that having determined the value of the property the cost of any given service was a simple matter. As a matter of fact, however, it is exceedingly difficult to determine the cost of any utility service or any freight shipment. How much of the cost of the roadbed, the locomotive, the fuel, the wages of the train crew, the terminal facilities, the salaries of the officials and office force, and the like should be allocated to any particular shipment? If rates are somewhat in proportion to distance, it is clear that in so far such rates are at least a partial application of the cost-of-service principle. Thus the freight rates in the Eastern territory are based on the New York-Chicago rate and are proportional to distance.

On the other hand, in the Southern and Western territories the distance principle is to a large extent violated. In the South, before 1916, many important points known as basing points were given exceptionally favorable rates. Thus the rate from New York to X, located in the vicinity of Atlanta, was the New York to Atlanta rate plus the rate from Atlanta to X, even though the distance from New York to X were less than the distance from New York to Atlanta. Since 1916 blanket rates for certain areas have been introduced. Rates from the East to intermountain towns were formerly higher than the rates to Pacific-coast towns because of the water competition by way of the Panama Canal. In recent years these rates have been equalized. Similarly, all places east of the Mississippi River pay the same rate to the Pacific coast and intermountain towns regardless of the distance. Pacific-coast commodities, such as fruit, lumber, wool, and fish, pay a blanket rate ("postage-stamp system") to all points east of the Rocky Mountains.

Moreover, the cost principle is violated as to different commodities. The rates on high-grade and expensive commodities, such as cutlery or jewelry, are much higher per pound or per unit of space occupied than are the rates on sand, coal, or iron ore. Such differentiation cannot be justified on the basis of the cost-of-service principle. The principle applied is rather the one common to monopolies — what the traffic will bear. Expensive commodities are charged a higher rate than are cheap, bulky commodities, because a higher rate can be charged without losing the shipment; whereas if a higher rate were charged on coal, iron ore, or sand, the commodity could probably not be shipped at all with profit to the producers. Thus the structure of rates on American railroads is largely an outgrowth of the policy of charging what the traffic will bear.

It should be noted that charging what the traffic will bear often means an exceptionally low charge, but it will never fall permanently below the *additional* cost incurred in hauling the commodity. If a commodity cannot contribute something toward the general expenses, it will not be worth while to haul it at all. On the other hand, expensive articles will stand a charge far in excess of their proportional share of the total expense.

Without public regulation a railroad will naturally tend to charge for each commodity a rate that will yield the maximum net income. Under public regulation the commodities that can stand only a small charge will not be materially affected. The commodities that will stand a heavy charge will be reduced so that the general level of rates will not yield a net revenue in excess of a fair return on a fair value of the property. Strictly speaking, public regulation may result in the reduction of any rate which is above the variable cost; that is, the added cost incurred in handling the commodity. On the other hand, many rates must be left considerably above the total average cost of handling the commodity in order to make up for the rates that fall below the total average cost for certain commodities.

How Railroad Rates are Actually Made. The present railroad-rate structure is a product of historical growth. It has been established, in the main, by the railroads themselves. The railroads must, however, file their schedule of rates with the Interstate Commerce Commission, and they cannot depart from these published rates except after thirty days' notice. But the Interstate Commerce Commission has the power to suspend any proposed changes in rates for a total period of five months, and upon examination it may set aside the rates filed by the railroads and substitute a new rate. The railroads may appeal to the courts for a review of the rates established by the Commission. If the rates can be shown to be so low that they amount to a confiscation of railroad property, the courts will set them aside as unconstitutional, because such rates amount to a taking of property without compensation. Thus the railroads make the rates in the first instance; the Interstate Commerce Commission has the power to review them, to set them aside if they are unreasonable, and to establish new rates. The rates thus established may be set aside by the courts if they can be shown to be confiscatory.

The Transportation Act of 1920. The Transportation Act of 1920 directed the Commission to adjust rates so that the carriers as a whole or in each rate territory may earn a fair return on the fair value of the property. The Commission in 1922 held that $5\frac{3}{4}$ per cent should be considered a fair rate of return. This does not mean, however, that the government guarantees the railroads a rate of $5\frac{3}{4}$ per cent on a fair value of their property: the government does not engage to make up any deficit. It is intended that the Commission shall allow sufficiently high rates to permit an average return of $5\frac{3}{4}$ per cent on common-carrier property. But if any or all of the railroads fail in any one year to make this $5\frac{3}{4}$ per cent, the government is under no obligation to make up the deficit. Nor is it expected that each individual railroad will be able to earn the $5\frac{3}{4}$ per cent: some will average less; others will average more. It is only anticipated that rates should be sufficiently high so that the average for all the roads may equal $5\frac{3}{4}$ per cent, and even this average rate is only aimed at but not assured. In fact, in 1920 the railroads of Class I earned only 0.09 per cent on their investment; in 1921 they earned only 2.91 per cent. In 1922, 1923, 1924, 1925, 1926, and 1927 the earnings were 3.6, 4.4, 4.3, 4.9, 5.1, and 4.5 per cent respectively. Moreover, if any individual railroad should make more than 6 per cent in any one year, the government is entitled to take half the excess, and out of the fund so derived to assist lines that are necessary, but unable to make earnings.

Federal Control over Security Issues. The Act of 1920 recognizes for the first time the importance of Federal control over the issue of securities. Although railroads should be permitted to earn a fair return on a *fair value only,* regardless of the outstanding issue of securities, yet the fact remains that a railroad's policies are influenced by a desire to pay dividends on its stock even though the capitalization does not correspond to the fair value. The Transportation Act of 1920 therefore grants exclusive power to the Commission to regulate the issue of securities of interstate railroads. Moreover, the Commission is granted control over the manner in which the proceeds are employed.

Railroad Consolidation. The Act of 1920 contemplates the ultimate consolidation of the railroads into a few great systems. There are many weak roads that find it difficult, if not impossible, to earn a fair return on their property. When rate adjustments are in progress these weak roads, with their insufficient earnings, afford opportunity for an argument in favor of rate advances, even though the stronger roads are earning a good return on their property. If certain weak roads were consolidated with certain strong roads into single systems, the problem of adjusting rates to a fair return on the fair value of the entire property would be greatly simplified. Under the 1920 act the Commission was directed to prepare a plan for consolidation, but any plan adopted by it was not made compulsory on the railroads. Under the direction of the Commission, Professor William Z. Ripley prepared a plan for the complete consolidation of the entire railroad structure into nineteen huge systems. Hearings on this plan have been held as required by statute. Competition must, according to the act, be preserved as fully as possible. It is therefore impossible to consolidate all the railroads in any one geographical area. Moreover, the consolidations are to be made in such a way that the various systems, while charging the same rates, may nevertheless be able to earn substantially the same rate of return on the fair value of their respective properties. There are many difficulties in the way; and it is not likely that voluntary consolidation, such as the act contemplates, will in the end prove successful.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The following table[1] indicates the development of railroads in the United States from 1890 to 1920. Compare with the growth of population.

	PROPERTY INVESTED (IN MILLIONS)	TONS OF FREIGHT CARRIED 1 MILE (IN MILLIONS)	PASSENGERS CARRIED 1 MILE (IN MILLIONS)
1890	$ 7.8	76	12
1920	19.8	410	47

---

[1] See Report of the Joint Commission of Agricultural Inquiry, Part III. Data given in the following problems are also taken from this report.

2. The railroads purchased 3,600,000 tons of rails in 1906, 2,500,000 tons in 1920, and 2,200,000 tons in 1921. What conclusions do you draw from these figures?

3. The following table gives (in thousands) the number of carloads of different commodities loaded in 1921:

Grain and grain products	2,281
Live stock	1,495
Coal	7,934
Forest products	2,483
Ore	904
Merchandise	10,677

4. The train revenue from freight service amounted in 1920 to $4,300,000,000, and the passenger revenue to $1,300,000,000.

5. From 1916 to 1920 the trend of railroad stocks was downward and the trend of industrial stocks was upward. How can you account for this?

6. The replacement value of all graded and surfaced roads in the United States in 1921 was $1,800,000,000. Compare this with railroad property.

7. In Canada there are two main railroad systems: (1) the Canadian National Railways, composed of those lines which have been unable to finance themselves and which the government has been forced to take over in order to maintain the service; (2) the Canadian Pacific Railway Company, privately owned and operated. What considerations would have to be borne in mind in comparing the government system with the private system?

CHAPTER XIX

SPECULATION AND PRICES

Production for the Future. The production and sale of goods under modern industrial conditions and for modern markets, which sometimes embrace the entire world, always involve more or less uncertainty. Cotton manufacturers who produce for the domestic or the foreign market cannot foresee exactly the price at which they will sell their cloth; the merchant who lays in a stock of goods cannot tell in advance the prices at which his patrons will buy; and the farmer cannot know in advance the demand which will exist for his products when they are finally brought to the market place. Inability to forecast demand is thus seen to be one of the prime causes of business uncertainty. And this uncertainty is partly due to the widening of the market. When production was carried on to supply a local, restricted demand, uncertainty was far less. It requires the exercise of much higher powers of business judgment than were necessary when most enterprises drew their raw materials from sources near at hand and sold their products to consumers living in a relatively small geographical area. Another cause is the fact that capitalistic production, involving as it does the use of much fixed, durable capital, necessitates planning ahead for relatively long periods of time.

Not only is the demand for goods uncertain, but the total market supply is likewise not definitely predictable. In agriculture and, to a less extent, in other extractive industries the annual product is dependent upon natural conditions, such as the weather, which are beyond man's control. Nor is the capricious behavior of nature the only source of uncertainty in production. Workers may strike without warning; credit may be expanded or contracted owing to a change in the policy of the banks; business may become dull under conditions that at another time seem not to affect in the least the upward trend of output, simply because at one time a majority of people were in a cautious mood, whereas at another time they were sanguine.

All the types of uncertainty which have been enumerated are purely "business risks." But these do not make up the entire list of uncertainties which surround the conduct of a large enterprise. There are risks due to fire, storm, and other natural hazards, which impinge on the community generally, including the business enterprise. But although these hazards affect prices, they are to a large extent insurable and can be transferred by the business man to insurance companies. The cost of replacing destroyed property is spread over the entire industry or community, each business or each locality paying its proportionate share for the sake of escaping a less certain but very much greater possible loss.

Methods of Meeting Business Risks. If future events could be accurately predicted, entrepreneurs would be able to calculate the future costs of making good their losses, and these losses would be added to costs of production. Prices would be affected by a change in the magnitude of the losses, to be sure, but there would be no business risks. As things are, it is not the total amount of the losses that embarrasses the business man but the impossibility of being prepared to meet them.

One obvious method of eliminating uncertainty, and therefore risk, is the perfection of knowledge of future events. Just to the extent that the cotton manufacturer knows in advance that the foreign demand for his products will be small, he avoids loss from overproduction for that market. And if the merchant ascertains in advance that sales in his territory are certain to be reduced during the coming season because of dullness in other lines of business, he can escape the loss of having to sell a part of his stock at less than cost. All methods of increasing the amount of information available for the use of business in forecasting future demand and future supply must tend to reduce risks. In the United States there are four types of services that tend to reduce the uncertainty of business: (1) the Federal government collects and publishes a great deal of current data concerning the acreage, condition, and yield of crops, the general conditions of business in foreign countries, and the trend of foreign markets.[1] It also publishes information as to the output of factories and mines, prices, and financial data at short intervals, and an annual summary of the same sort of data together with information on power used by producers, the number of workers employed during the year, and the value of the products of important industries.[2] (2) Certain informational services are maintained by private concerns and sold to subscribers. Some of these services are general in nature; others pertain to particular trades or groups of trades. (3) Trade and financial journals collect and disseminate information partly in the form of descriptive statements to the effect that sales are "poor" or "good" in different localities or different trades, or that collections are "slow" or "good." Much statistically descriptive information is also included in these statements. (4) Some large concerns maintain their own informational services. Facts are collected, in quantitative form when possible, and analyzed, and conclusions are drawn for the use of the managers.

When any concern can obtain more accurate information about future conditions than its competitors, it can make greater net gains and, as far as competition is complete, drive them out of business. By avoiding losses it can reduce costs of production; and by knowing when and how much to produce, it can better serve its clientele. Society also gains from the elimination of the losses incident to business uncertainty; for if one competitor can reduce risks by perfecting its information concerning future markets, its competitors will imitate its methods, costs to marginal producers will fall, the consumers will buy at lower prices, and the national income will be increased.

---

[1] See *Crops and Markets* and *The Agricultural Situation.*

[2] See the *Survey of Current Business* and the *Commerce Yearbook.*

Still another method of reducing uncertainty is by consolidation. A producer who sells his wares over a limited geographical area must, in the absence of foreknowledge of the variations of demand and supply in that market, share all the ups and downs of general business in that area. But some of these variations can be escaped by selling over a larger area. Many of the causes of variations in demand and supply are local in origin and effect. Partial crop failures and overproduction are often limited to a small region. A merchandising concern that has many stores scattered throughout a country, or a packing company that sells meat over many districts differing in resources and industries, can often take advantage of its wide distribution by selling in other districts goods which would normally be demanded in one district but which cannot be sold there because of some unforeseen disturbance of local incomes; or, if competition becomes so severe that costs cannot be covered in one place, it can withdraw or temporarily cut down its supply and dispose of its products in places where the reverse conditions prevail.

When many risks are thus consolidated or brought together and assumed by a single concern, there is a gain in predictability. The company which insures dwellings against fire can foretell very little better than the owner the chances that a single house will burn down in the course of a year; but it can predict with much greater certainty how many houses out of several hundred thousand scattered over a wide area will be destroyed. It can do this because the chance conditions which cause fires follow a normal course through the years, and a greater prevalence in one locality is offset by a lesser prevalence in another. In the same way the large business concern that can distribute its sales over a wide area or that can produce several noncompeting commodities can rely on the constancy of phenomena when a large number of cases are dealt with.

Speculation and Risk-Taking. When future events are uncertain, different people are likely to have divergent opinions about them. This is true of future prices. Vacant parcels of land in a growing city are almost certain to be used at some time in the future; but the exact time at which they will be improved and the value they will have at that time are uncertain, and among a group of real-estate dealers who know the local conditions there will almost certainly be some difference in estimates in respect to these two important future events. The present owners of these tracts of land may believe that future values warrant a present price of $5000 a lot. If other persons believe that this price is too low, they will probably buy the lots and wait for the future rise in values. In doing this they are engaging in speculation. This is the simplest and most ancient type of speculative dealing: buying a good outright in the expectation of making a gain through a rise in price that is greater than the rise expected by its owners. Speculative deals of this sort are, of course, exceedingly common; but usually they do not attract much attention because they are mostly confined to local markets, because their volume is unknown, and because they have the sanction of long-established usage.

The speculation about which the public is concerned today is that which takes place in the stock and produce exchanges. These are large markets; and because they are large the total volume of sales attracts the attention of the entire country. They are not, of course, confined to speculative operations. An investor who wishes either to buy or to sell stocks and bonds finds in the New York Stock Exchange a continuous market for listed securities. The owner of a country elevator or the farmer may sell his wheat on the Chicago Board of Trade, and the miller who wants wheat to supply his mills may buy in the same market. But by far the larger portions of all sales in these two types of market are speculative deals.

In the produce markets the speculation takes the form of "dealing in futures," which is distinguished from "cash sales."

In the latter, grains and other agricultural products are bought from samples and the goods are actually delivered to the buyer; in the deal in futures the seller merely enters into a contract to deliver to the buyer (say) wheat of a certain grade at some future date. Such a bargain takes place just as it does in the real-estate market — because two people happen to have different estimates as to the probable future price. The seller believes that May wheat should sell somewhat below $1.50 a bushel; the buyer believes that it should sell above that price. Since all those who believe that the price will be less than $1.50 want the market to go lower, they are called "bears," and their operations are said to *bear* the market; those who believe the market will go higher and want it to go higher are called "bulls," and they are said to *bull* the market. In a typical deal in May futures the buyer enters into a contract with the seller to take from him in May, say, 10,000 bushels of wheat at $1.50 a bushel. If the seller can go into the cash market and buy wheat of the stipulated grade for $1.40 a bushel, when the time for fulfilling the contract approaches he will make 10 cents a bushel on the entire lot and the buyer will lose the same amount of money. But if the seller can cover his contract only by buying at $1.60 a bushel, he will lose and the buyer will gain.

As a rule no wheat is delivered on such contracts. The seller of futures will instead buy up a contract to buy and turn it in to liquidate his obligation. Both the buyer and the seller may change their positions several times before May arrives, buying at one time and selling at another. The net effect of such transactions where actual delivery is not even contemplated, is simply that whoever finds that his contracts, if carried out, would bring him a loss must pay that amount to others whose contracts would entitle them to a gain.

On the stock exchange speculative transactions have the same general features, with the difference that future contracts are not entered into, being prohibited by the rules of the exchange. Instead, the seller goes to the loan counter and "borrows" the stock in order to make actual physical delivery. He must deposit a sum sufficient for the security of the borrowed stock at its market price. The deals take place for the same reasons as on the grain exchanges. The seller believes that a given stock is selling at a price that the future prospects of the issuing company or general business conditions do not warrant. He believes that the stock will go down in price; the buyer believes that it will rise. If the stock goes down then the seller can "cover" (that is, replace the shares borrowed) at a lower figure, collect the difference between the price at the time he sold and the (then) ruling price, and pocket the difference; but the buyer loses. If the stock rises in price, then the buyer gains and the seller loses.

Speculative transactions in both grain and stocks exhibit strong points of resemblance to gambling. There are differences however. The chance on which the gambler bets is created by him for the purpose of gambling. The chances, or uncertainties, on which the traders in the exchanges risk their money are not of their making. As we have seen, they are the unavoidable concomitants of modern business methods. They persist because men are unable to forecast changes in demand and supply. In the second place, the gambler does not even incidentally contribute to the production of goods, whereas the speculator may do so.

Effect of Speculation on Prices. When a speculator in lands buys in the expectation of a greater rise in value than is commonly anticipated in the market, he raises the demand for such land and he also tends to raise the price of similarly situated parcels. If he is inexperienced he may, of course, be entirely mistaken, and his estimate will have little effect upon the market unless he is only one of many such amateur traders. In the latter case a boom may develop which will benefit no one but the original owners of the property. The amateurs will lose money and the community will be poorer by the amount of time and money wasted in carrying through the deals. On the other hand, if the speculators are experts it is likely that their guesses as to the future values of the land will be better than those of the owners whose experience and training have been in some other occupation. In that case there may easily be a social gain. For example, the owner of a tract of land may contemplate selling it to a warehouse company. But a speculator who believes that within a few years the land will be more valuable for retail shops outbids the warehouse company and holds the land for the more valuable use. If this forecast of the future development of the city proves correct, then both he and the community will have gained as a result of the speculation.

In a speculation in wheat a buyer may believe that the November price for wheat deliverable in May is too low. His knowledge of the market leads him to believe that the crop forecasts in the Argentine are too optimistic or that the demand will be heavier than has been supposed. He concludes that by May the price will be higher than it now is on future contracts, and so he buys for delivery at that time. If his guess proves correct, he will make a profit. But has his action been of any benefit to the community? In the first place, the fact that he has bought has increased the demand for May delivery; and if his purchases are large enough, the price will be forced up on that account. And the rise in the May price will affect the cash price in November. The result will be a more equal distribution of the wheat for use during the year and a more nearly uniform price throughout the year. To be sure, these obviously desirable results are realized only if the speculators tend to make more correct estimates than other people. But the bulk of the trading on any large grain exchange is done by experts, and the estimates of future demand and supply which they make are much more likely to be correct than are the guesses of outsiders or even of farmers and local grain-buyers. The existence of the organized speculation thus does result in some social gain.

There is only one way to avoid the risk of price changes, and that is by consuming as soon as production has been completed. The farmer who harvests a thousand bushels of wheat must sell at once if he is to avoid risk. If he holds his grain, it may either rise or fall in price. But not all farmers can sell their crops as soon as harvested. Even if they did, the crops would not be immediately consumed, and the dealer who held them would be compelled to assume the risk. It is the primary business of a farmer to grow grain and of the local buyer to gather it together into carloads, clean it, and dispatch it to a central market. Even the warehouse company that stores it at the central market has a special occupation, and so has the miller who grinds it into flour. But the risk of price changes exists, and in the absence of a specialized body of risk-takers, it falls on the person who holds the grain. It is the function of the traders on the exchange to assume the risks — or most of them — of price changes. Since they are experts it may be supposed that they will make fewer mistakes than any of the other persons concerned in the production and distribution of grain and its products.

The risk of variations in price can be transferred to the specialists in risk-taking by the simple device of a hedge. A farmer whose wheat is in the granary by September may be unable to haul it to market until three months later. If he wishes to avoid the risk of changing prices, he can sell December wheat in the futures market and leave his wheat in the granary until it is convenient to deliver it to the local elevator. If the price falls lower than he had anticipated, he will be able to cover the loss on his crop by the gain on his transaction in the futures. The manager of a local elevator may cover the risk of price change while his shipments are in transit by selling futures in the exchange. The miller who has filled his warehouses with grain stands the risk of losing money if the price of wheat (and consequently of flour) falls before he has milled and sold the wheat. He too may escape risk — transfer it to the traders — by selling wheat for future delivery.

It must not be inferred that all the dealing on the grain exchanges contributes to desirable ends. When risks are assumed by those most competent to take them, and when the dealing in the exchanges causes prices to fluctuate less than would otherwise be the case, there is a social gain. The more even distribution of consumption in accordance with periodic supplies is also a social gain. For if too great consumption at abnormally low prices occurs at one time, it must be offset by diminished consumption at higher prices at a later time. We know from our study of the principle of diminishing utility that the satisfaction consumers derive from an additional block of goods does not offset the deprivation in utility suffered from a shortage later on. As the amount consumed is diminished, the total satisfaction enjoyed by the users diminishes more than proportionally; and as the amount increases, the total satisfaction increases less than proportionally.

Speculation in the markets for securities stands on somewhat different grounds from that in the grain exchanges as far as the social effects are concerned. The things dealt in are not commodities but claims to income and property rights that give control of producing plants; and while the most efficient distribution of the investment capital of the country is undoubtedly facilitated by the existence of an organized market for the buying and selling of stocks and bonds, speculation in them is so largely a matter of betting on the probable future earnings of the corporations whose securities are dealt in, or on the future cost of obtaining bank loans for speculative operations, or even on fluctuations resulting from speculative buying and selling, that the social gain is doubtful.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The man who does not insure his house against fire is a speculator; the man who insures is not." Are these statements true?

2. The sales of futures on the Chicago Grain Exchange are several times as great as the cash sales. The men who trade in futures deal in "phantom" wheat, and the supply of this phantom wheat must cause the price of real wheat to be lower than it would otherwise be. Is the foregoing statement correct?

3. The variation in the prices of agricultural products not dealt in on the great speculative exchanges is greater in normal times than is the variation in the prices of grains which are bought and sold on these exchanges. Why should this be true?

4. "During the boom in land prices in Iowa in 1919–1920 there was much speculation. Lands were greatly overvalued, and when the decline in prices occurred many farmers were forced into bankruptcy. This was a great social loss." Argue for or against the statement "this was a great social loss." Was the extravagant rise in prices due to speculation or to something else?

5. "Men speculate because prices fluctuate; prices do not fluctuate because men speculate." Is this statement correct? Explain.

6. If a factory-owner buys a new machine he cannot be sure that his investment will prove profitable. Is his investment a speculation? Does he take a risk?

BOOK III. MONEY AND PRICES

CHAPTER XX

MONEY AND CREDIT

The Functions of Money. In the foregoing chapters we have discussed value and price. In modern societies the value of any commodity is measured in terms of money. Value so expressed is called price. In other words, *price* is the quantity of money for which a commodity will exchange in the market. Money, therefore, is the yardstick by which we measure value. But unlike other units of measurement, such as the yard or the pound, the monetary unit is itself a fluctuating one. The yard is the unit by which we measure space, and it is of constant length. The pound is the unit by which we measure weight, and the pound is of constant weight. The dollar is the unit by which we measure value, but, strangely enough, it is of fluctuating value. Thus the price of a bushel of wheat may double, but the value of wheat may be exactly the same as before. In this case the unit by which we measure value has itself changed in value. Prices must thus be distinguished sharply from exchange values. Likewise, the demand for a commodity in terms of money may have greatly increased, yet the real demand in terms of other commodities may not have changed at all.

Money not only serves as the yardstick by which value is measured in modern society; it also serves as a medium of exchange. Goods and services are not exchanged against goods and services: they are exchanged for money or its equivalent. These, then, are two of the important functions which money performs: first, it serves as a measure of value; second, it serves as a medium of exchange. In addition, as will be noted in a later section in this chapter, certain kinds of money perform a special function in the capacity of government and bank reserves.

Money and Circulating Media. "Money" is a broad term; it includes a great many varieties and species. Not all forms of money serve as the measure of value; only standard money performs that function. On the other hand, there is in all modern countries a great quantity of circulating media which are not, properly speaking, money. We must therefore distinguish between *standard money*, *money*, and *circulating media.* It would, in fact, be more accurate to say that standard money serves as the measure of value and that money and money substitutes serve as the media of exchange.

Money is any medium of exchange which passes freely between persons *without indorsement*; or, in other words, money is any medium of exchange which is acceptable on its own account, regardless of the financial standing or integrity of the person who offers it in exchange for goods or services. Checks, drafts, money orders, etc. are forms of circulating media which are not acceptable on their own account but which will readily be accepted if indorsed by someone favorably known to the person to whom they are presented.

Kinds of Money in the United States. Standard money is that form of money in which all other forms of money are redeemable. Gold coin is the standard money in the United States. The gold dollar, if coined, would contain 23.22 grains of pure gold. The gold dollar is, as a matter of fact, not coined; but the eagle, or ten-dollar gold piece, contains 232.2 grains of pure gold. All other forms of money are finally redeemable in gold. Although silver dollars are not directly redeemable in gold, they ultimately would be if necessary, since the Secretary of the Treasury is by law intrusted with the duty of maintaining all other forms of money at parity with gold. To accomplish this result the government must, if necessary, stand ready to exchange gold coins for silver dollars.

The following is an outline of the different kinds of money in circulation in the United States:

A. Government money, or "lawful money," so called
  1. Specie
    a. Gold
    b. Silver
    c. Subsidiary coin
  2. Paper
    a. Gold certificates
    b. Silver certificates
    c. United States notes (greenbacks)
    d. Treasury notes of 1890
B. Bank money
  1. National bank notes
  2. Federal Reserve bank notes
  3. Federal Reserve notes

Gold is coined into quarter-eagles ($2.50), half-eagles ($5), eagles ($10), double-eagles ($20). Besides the silver dollars there are the subsidiary coins: the half-dollars, quarters, dimes, nickels, and pennies. Gold and silver certificates may be described as virtually equivalent to warehouse receipts for gold or silver. Most people prefer paper money to the clumsy silver dollars or the easily lost gold pieces; hence gold and silver circulate to a large extent in the form of gold and silver certificates, the gold or silver being deposited in the vaults of United States Treasury as security for the outstanding certificates.

United States Treasury notes, or "greenbacks," amounting to $450,000,000 were issued during the Civil War. Some were retired following the war, but the total volume was finally fixed by law at $346,681,000. They are promises on the part of the United States government to pay to the bearer on demand. During the greenback period (1862–1879) the government was unable to redeem this promise, but since that date greenbacks have been redeemable in gold. The government has for many years maintained a gold reserve of $150,000,000 in the Treasury to guarantee redemption.

The treasury notes of 1890, once of considerable importance, have almost disappeared. They were issued under the Sherman Silver Purchase Act, which provided that the Secretary of the Treasury was to purchase 4,500,000 ounces of silver per month. This silver was paid for by the issuance of treasury notes. The silver bullion thus purchased was the security for the notes. Therefore these treasury notes are virtually silver circulating in the form of paper money.

The National bank notes are issued by national banks. They are obligations of the bank issuing them, and are secured by government bonds of equal value, deposited with the Comptroller of the Currency at Washington, besides a deposit of lawful money equivalent to 5 per cent of the notes issued. These notes are, moreover, a first lien on all the assets of the issuing bank. Federal Reserve bank notes are identical with National bank notes, except that they are issued by Federal Reserve Banks. Under the old national banking system, before the establishment of the Federal Reserve system, national banks were virtually compelled to issue notes, since they were required to invest a portion of their capital in government bonds. They may now dispose of these bonds up to $25,000,000 per year, in which case the bonds are absorbed by the Federal Reserve Banks, and on the basis of these bonds Federal Reserve bank notes may be issued.

Federal Reserve notes are issued by the twelve Federal Reserve Banks. How they are secured will be described elsewhere in this chapter.

Table XII shows the quantity of the different kinds of money obtaining in the United States on January 1, 1928. The table gives the total stock of the different kinds of money, the amount held by the Treasury, the amount held by the twelve Federal Reserve Banks, and the amount "in circulation," by which is here meant the amount in the hands of banks (other than the Federal Reserve Banks) and the population generally.

If we subtract from the money held in the Treasury the gold and silver which is held in trust against gold and silver certificates, there remains roughly $1,921,000,000. If we sub-

Table XII

Kind of Money	Total Stock in United States	Held in the Treasury	Held by Federal Reserve Banks and Agents	In Circulation
Gold coin and bullion	$4,376,258,102	$3,503,422,503[2]	$473,758,088	$399,087,511
Gold certificates	(1,617,019,369)[1]		543,284,480	1,073,734,889
Standard silver dollars	537,970,158	477,323,017[3]	11,872,387	48,774,754
Silver certificates	(472,529,901)[1]		72,261,105	400,268,796
Treasury notes of 1890	(1,314,850)[1]			1,314,850
Subsidiary silver	300,735,507	2,143,049	12,990,131	285,602,327
Minor coin	115,250,575	1,384,146	2,095,913	111,770,516
United States notes	346,681,016	3,962,625	49,829,716	292,888,675
Federal Reserve notes	2,234,096,545	3,568,870	467,734,124	1,762,793,551
Federal Reserve bank notes	4,439,688	122,408	35,682	4,281,598
National bank notes	701,003,589	19,940,364	61,624,630	619,438,595
Total	$8,616,435,180	$4,011,866,982[4]	$1,695,486,256	$4,999,956,062

tract the $155,000,000 held against the greenbacks and the treasury notes of 1890, there remains $1,766,000,000 in the Treasury, of which $1,556,000,000 was held for Federal Reserve Banks and a further $32,000,000 (approximately) was held for redemption of National bank and Federal Reserve bank notes. The government therefore had about $175,000,000 available for its own use. If we add to the money held by the Federal

---

[1] These amounts are not included in the total, since the money held in trust against gold and silver certificates and treasury notes of 1890 is included under gold coin and bullion and standard silver dollars respectively.

[2] Of this amount $1,556,510,011 was held for Federal Reserve Banks and agents, $1,617,019,369 was held in trust against gold certificates, and $155,420,721 was held as reserve against United States notes and treasury notes of 1890.

[3] Of this amount $473,844,751 was held in trust against silver certificates.

[4] The amount of money held in trust against gold and silver certificates and treasury notes of 1890 should be deducted from this total before combining it with total money outside the Treasury to arrive at the stock of money in the United States.

Reserve Banks the $1,556,000,000 held for them by the Treasury, we find that the total money controlled by the Federal Reserve Banks was about $3,252,000,000. Of the $5,000,000,000 in circulation, so called, a considerable portion was held in the thousands of state and national banks.

The Standard of Value. In 1928 we had in the United States about $4,500,000,000 of gold. Gold is our standard money. It is our measure of value. But gold has not always been our standard money. From 1792 (when our first coinage act was passed) to 1834 silver was the actual standard; from 1834 to 1862 gold was the actual standard; from 1862 to 1879 greenbacks were the actual standard; and from 1879 to the present gold has again been the standard of value.

This brief survey of our standard money leads us to a consideration of our monetary history. In 1792 we started out with a double standard, both gold and silver being by law standard money. When gold and silver are both made standard money by law, we have what is known as bimetallism. Bimetallism exists when both metals are coined freely; that is, when there is no limit to the quantity of either metal which may be coined into money, any person being free to bring as much gold or silver bullion as he wishes to the mint to have it coined into money. Both metals when coined into money must also be legal tender; that is, both must be accepted by law in payment of a debt unless the contract specifies otherwise. To put it another way, if both gold and silver are legal tender, any debtor may legally discharge his debt by offering to the creditor either gold or silver money. Free (that is, unlimited) coinage of both metals and equal legal-tender qualities are the two conditions necessary for a bimetallic money standard.

In 1792 our silver dollars contained 371.25 grains of pure silver, and all through our history this has been the weight of our standard silver dollars. In 1792 the gold dollar contained 24.75 grains of pure gold. It will thus be evident that the weight of the silver dollar was fifteen times the weight of the gold dollar. The mint ratio was therefore 15 to 1. This ratio was adopted because on the market one grain of gold bullion sold for approximately fifteen grains of silver bullion. Thus the pure silver contained in the silver dollar was worth approximately the same as the pure gold in the gold dollar. The market ratio was approximately 15 to 1.

But soon the market ratio changed to approximately $15\frac{1}{2}$ to 1. Since the mint ratio was fixed by law at 15 to 1, it is clear that there was not enough silver in the silver dollar to equal in value the gold in the gold dollar. Under these circumstances it is clear that no one would pay his debts in gold coin. The gold content of the gold dollar (24.75 grains of gold) would now buy 383.62 grains of silver in the bullion market; and since it required only 371.25 grains to make a silver dollar, one would always be paying the lesser value by paying one's debts in silver dollars. The gold dollar, as a dollar, was worth no more by law than the silver dollar, but the gold content of the gold dollar was worth more than the silver content of the silver dollar. Consequently gold was worth more as bullion than as coin; and as a result it was no longer coined in practice, and coins already made were melted into bullion. Thus the "cheaper" of the two kinds of money, both standard by law, drove the "dearer" out of circulation. This tendency on the part of the cheaper money to displace dearer money is known in economic literature as Gresham's law (Gresham being the name of the man who first clearly stated it). While legally there were two standard forms of money, in actual fact silver became the standard of value.

In 1834 Congress changed the weight of the gold dollar with a view to bringing it back into circulation. The weight of the gold dollar was reduced to 23.22 grains; the silver dollar remained as before. This made the mint ratio approximately 16 to 1. The market ratio, however, was slightly lower, let us say $15\frac{1}{2}$ to 1. Under these circumstances no one would pay his debts in silver dollars. The silver content of the silver dollar (371.25 grains) would buy 23.95 grains of gold, whereas it required only 23.22 grains of gold to make a gold dollar. Silver was now more valuable as bullion than as coin. As a result silver was no longer coined into dollars, and existing silver dollars were soon melted down into bullion. Gold became the actual standard of value, and hence we may properly say that the silver content of the silver dollar was worth $1.03.

In order to keep subsidiary silver in circulation the weight of these coins was reduced in 1853. This accounts for the fact that two half-dollars weigh less than a silver dollar. Being coined in limited quantities, subsidiary silver was unable to drive out gold.

Gold then remained the actual standard of value until 1862, when, under the stress of war, the government issued huge quantities of paper money (greenbacks). These greenbacks were promises on the part of the United States government to pay on demand. But by reason of the excessive issue the promise could not be redeemed; so in relation to gold the greenbacks fell in value. Since greenbacks (though irredeemable) were nevertheless legal tender, everyone naturally preferred to pay his debts with them rather than with gold. The result was that gold dollars disappeared from circulation (except in California) and greenbacks became the actual standard of value.

After 1865 the greenbacks gradually rose in value, and finally (1879) they became redeemable in gold. Thus again gold became the actual standard of value, and it has remained our standard from 1879 to the present time.

There was in the nineties, however, a period in which the gold standard was again seriously threatened. In 1876 rich silver mines were opened up in Colorado. Silver became cheap compared with gold. The market ratio rose rapidly above the 16 to 1 mint ratio until in 1878 it was 18 to 1. It then became profitable to coin silver into dollars; but the law no longer permitted free coinage of silver, for in 1873 an act had been passed which demonetized silver. This had probably been done without any malice aforethought. Silver dollars had automatically disappeared from circulation under the operation of Gresham's law as a result of the act of 1834, which established the mint ratio of 16 to 1. When silver became cheaper and the market ratio rose to 17 or 18 to 1, people who had silver bullion (such as the silver-mine operators) naturally wished to be able to convert it into dollars at the mint ratio of 16 to 1. A silver dollar would be as good as a gold dollar in paying off a debt, but the silver content of the silver dollar was not worth the gold content of the gold dollar. Furthermore, in the period of the seventies the supply of money failed to keep pace with the rapidly expanding volume of trade, prices were falling (as will be explained later), and therefore many people who were suffering from the fall in prices favored a return to bimetallism. Thus arose the silver movement in politics. Congress passed in 1878 the Bland-Allison Act, which required the Treasury Department to purchase silver bullion and coin it into silver dollars of legal tender. But the amount of silver that could be coined under this law in any one month was limited; in fact, only about 25,000,000 silver dollars were coined each year from 1878 to 1890. These silver dollars were placed in the Treasury, and silver certificates were issued redeemable in the dollars so deposited.

This act did not satisfy the demands of the growing silver movement. As a result the Sherman Silver Purchase Act was passed in 1890. This act required the Secretary of the Treasury to purchase 4,500,000 ounces of silver per month, the silver to be paid for by the issue of treasury notes. These were the treasury notes of 1890.

Some 576,000,000 silver dollars were coined from 1878 to 1893 as a result of these two acts. The value of the silver content of these dollars was less than the value of the gold content of a gold dollar; therefore people naturally wanted to exchange these silver dollars for gold dollars if possible. As the treasury notes were by law redeemable at the Treasury Department in either gold or silver, the silver dollars coined under the Sherman Act were virtually redeemable at the Treasury Department in gold, since the treasury notes which were based on the silver *were* redeemable in either gold or silver. In practice, therefore, treasury notes were constantly being presented at the Treasury Department for gold. But although people were anxious to exchange silver dollars for gold dollars, they naturally preferred to pay their debts in silver, the cheaper metal. Gold was thus disappearing from the Treasury on the one hand and from circulation on the other. At this juncture Congress repealed the silver purchase clauses of the Sherman Act and saved the situation for the gold standard. Had the coinage of silver dollars continued, the silver dollar would soon have become the standard of value. The gold standard was established on a firmer legal basis by the Gold Standard Act of 1900, which provided that the Secretary of the Treasury should maintain at par with gold all forms of money issued or coined by the United States.

Thus ended a century of experimentation in the United States with the monetary device of bimetallism. There is only one reason for having a double standard — for having a standard dollar which may be either gold or silver. It is alleged, according to this reasoning, that when two metals are used there is less likelihood that the value of the standard dollar will fluctuate as much as when only one may be coined freely. Now everyone admits that the fluctuation of the standard of value is an evil; and if bimetallism would cure or even mitigate the difficulties arising from its fluctuations, we should all acclaim it superior to monometallism. But the reasoning rests on an assumption that has not stood the test of experience. It assumes that the natural flow of two metals from the mines will be more stable than the flow of one, but this has not proved to be the case. Silver is apparently more plentiful in nature than gold; it is found in larger deposits, and it is a by-product of the mining of copper and lead. Hence the scheme of bimetallism has worked in the past to produce greater fluctuations in the value of the dollar than if gold alone had been the sole source of standard money. There is a further difficulty, as we have shown by our sketch of the monetary history of the United States. It is not difficult to select a mint ratio which at any one time is approximately the same as the market ratio. But the market ratio depends upon the relative quantities of gold and silver coming from the mines. When this relationship changes, — when either gold or silver increases sharply, because of discoveries of bonanza mines, — then the market ratio changes. The mint ratio being fixed by law cannot change continuously, and the result is that the metal which is being produced more freely begins before long to drive the other out of circulation. This in itself is perhaps no evil; but there comes with this substitution of actual monometallism for theoretical bimetallism a rapid change in prices and a general uncertainty in respect to all payments of debts calling simply for "dollars."

Deposit Currency. Standard money is the measure of value, and all money is a medium of exchange. But in modern society money constitutes only a small part of our medium of exchange. In fact no more than 6 per cent of our trade is carried on by means of money; the other 94 per cent of our buying and selling operations is effected through money substitutes. The most important of these are the check, the bank draft, and the money order. Most purchases made in modern society are paid for by check. We have now to inquire into the origin and nature of these money substitutes.

The check is an order drawn upon a bank to pay money to the bearer of the check to whom it has been properly indorsed. The person who writes the check is the "drawer," or "maker," of the check, the bank upon which it is drawn is the "drawee," and the person to whom the money is to be paid is the "payee." Now it is obvious that a person cannot draw checks on a bank unless he has a deposit at the bank which entitles him to order the bank to pay out money.

This brings us to a consideration of the origin and nature of a deposit. The word implies that something has been deposited at the bank. You may go to your bank with a bag full of money and deposit it there, receiving a credit for the amount so deposited. This credit constitutes a "deposit account." It consists of a promise on the part of the bank to pay out money on demand for the amount of the deposit as you may order through the drawing of checks. Now it is clear that if all demand deposits originated in this fashion, they would virtually be warehouse receipts for money, and not in reality a money substitute at all. In fact, only a very small part of our demand deposits originates through the actual deposit of money.

At the close of a day's business a mail-order merchant may have received, in payment for goods sold, a bundle of checks, bank drafts, and money orders. He has probably received little or no money. He takes these checks, drafts, and money orders to his bank and deposits them. In a sense he may be said to sell them to the bank, and the bank pays him in return by crediting him with their value on his deposit account. The bank promises to pay money on demand in return for the checks, drafts, and money orders so deposited.

Credit Instruments. Most demand deposits do not arise, however, in either of the two ways referred to above. A manufacturer or a wholesaler sells an order of goods to a retailer. The retailer is unable to pay cash for it. He gives the wholesaler or the manufacturer his promissory note for sixty or ninety days. The manufacturer or the wholesaler takes this promissory note to the bank, indorses it, and "sells" it to the bank. The bank pays him the face value of the note minus the discount for sixty or ninety days. Does the bank pay him actual money? Not at all: it credits him on his deposit account. Thus demand deposits originate in large part through loans.

It may be that the manufacturer agreed to carry the retailer on his books for sixty or ninety days. But the manufacturer is in need of circulating media to pay his obligations. He therefore borrows at the bank on his own promissory note for sixty or ninety days. The bank discounts the note; that is, credits him on his deposit account with the face value of the note minus the discount.

There is still another way in which the sale to the retailer may be financed. The wholesaler, upon having sold the goods, may draw a draft on the retailer for sixty or ninety days. This draft, or bill of exchange, is an order drawn on the retailer to pay the sum of money in question. He then takes it to his bank and has it discounted. At the end of the sixty or ninety days the banker presents it to the retailer for payment.

Most deposits, then, do not originate through the deposit of money, checks, bank drafts, or money orders; they arise through the process of discounting. The bank "buys" promissory notes or bills of exchange and gives in exchange its own promise to pay money on demand. But the bank's promise to pay money on demand (bank credit) is, in practice, as good as money itself, since checks properly indorsed are usually accepted as media of exchange.

The Volume of Deposit Currency. Demand deposits are frequently referred to as "deposit currency." The total volume of deposit currency in the country at any one time may be said to be measured by the total volume of demand deposits. The volume of *checks* in circulation represents the volume of deposit currency *in use* at any given time. Only a small part of the total volume of deposit currency is in use at any one time, just as only a small part of the money in circulation is in use at any one time.

As a rule, every bank carries a deposit account with several other banks. These banks are known as its correspondents. Banks may draw upon these deposits just as individuals draw upon their checking accounts. They do so by writing bank drafts. If you owed a debt to somebody in New York to whom your check would not be acceptable, you would probably go to your bank and buy a bank draft and mail it to your creditor. Bank drafts are therefore really equivalent to checks, since both are drawn against deposit accounts. Deposit currency thus circulates in the form of checks or bank drafts.

Bank deposits, it will be remembered, are promises on the part of banks to pay money on demand. At first thought it might appear that the total volume of such deposits would never be large because they would soon be paid out in money on demand. It may therefore seem surprising that the volume of demand deposits is, year in and year out, approximately five times the quantity of money in circulation. Deposits are surely intended for use. Why are they not drawn upon? The answer is that they *are* constantly being drawn upon, but the checks are re-deposited as fast as they are drawn. Indeed, some checks are constantly being presented for actual money, but the money soon gets into the hands of someone who re-deposits it. Promissory notes and bills of exchange are constantly maturing, but others are constantly being discounted. It is clear that deposit currency constitutes a medium of exchange which is constantly in circulation side by side with money itself. To be sure, there are seasonal and cyclical fluctuations in the outstanding volume of deposit currency and money. At times relatively more of the total circulating media consists of deposit currency; at other times relatively more consists of money. Thus during the fall of the year in the crop-moving season there is a heavy demand for actual money. At such times deposits are being converted into money.

Banks keep far too little money in their vaults actually to pay the total volume of deposits if all deposit claims were presented for payment at any one time. They must, however, keep a certain amount of cash in reserve, since a small proportion of the deposits (as noted above) is constantly being demanded in money. This accounts for the fact that such a large proportion of the total money supply of the country is not in general circulation, but is held by the banks. About half the money of the country is so held.

We have now reached a point where it becomes necessary to describe briefly the structure of our banking institutions in order to see clearly the interrelations of cash reserves, money in general circulation, deposits, and loans.

The Federal Reserve System. At the head of our banking system are the twelve Federal Reserve Banks. To these are attached the member banks, consisting of all the national banks and many of the state banks and trust companies. The Federal Reserve Banks are bankers' banks. They do not receive deposits from anyone except member banks and the United States government. The member banks serve the public; the Federal Reserve Banks serve the member banks. The Federal Reserve Banks are required to keep a cash reserve of 35 per cent of lawful money against their deposits; the member banks are required to keep a reserve of 7, 10, or 13 per cent against their deposits, depending on which of three groups of cities the bank is located in. Minneapolis banks are required to keep 10 per cent. Some fifty cities are in the 10 per cent class. The vast majority are in the 7 per cent class, and only banks located in New York and Chicago are in the 13 per cent class. The member banks are required to keep all their legal reserves with the Federal Reserve Banks. Of course, they necessarily keep a considerable amount of cash on hand to meet their daily requirements. But the 7, 10, or 13 per cent reserve must be deposited with the Federal Reserve Bank in their district. These constitute the bulk of the deposits of the Federal Reserve Banks, against which, it will be remembered, only a 35 per cent cash reserve is required. Looking at the banking system as a whole, therefore, we see that only about $3.50[1] of actual cash need be available in the vaults of the banks to guarantee the redemption of every $100 of deposits owned by the public. As a rule, about twice that much is actually available.

Individual business firms, as we have seen, may present their promissory notes and bills of exchange for discount, and in return they ordinarily receive deposit credit. Member banks, in turn, may present this same commercial paper for rediscount to the Federal Reserve Bank in that district. When member banks re-discount their commercial paper in this manner with the Federal Reserve Bank, they ordinarily receive in return either *a deposit* at the Federal Reserve Bank or *Federal Reserve notes.* If the member bank wishes to increase its reserve in order to expand its loans, it will probably wish to take the deposit account; if, on the other hand, the member bank needs more money to pay out to its customers, it will be likely to take the Federal Reserve notes. It makes very little difference to the Federal Reserve Bank which is preferred. Federal Reserve notes are a promise of the Federal Reserve Bank issuing them, just as are also the deposits. The Federal Reserve Bank must maintain a gold reserve of 40 per cent against these notes and, in addition, a sufficient amount of commercial paper to equal, together with the gold reserves, not less than 100 per cent of the notes issued. It is therefore slightly more advantageous for the Federal Reserve Bank to credit the member bank with a deposit, since a reserve of only 35 per cent is required in the case of deposits, and the reserve may be in any form of lawful money. Since, however, the vast bulk of the lawful money held by the Federal Reserve Banks is in gold anyway, the difference between these requirements is not great. It is therefore clear that the Federal Reserve notes are closely analogous to deposit accounts with the Federal Reserve Banks. But only member banks can have deposits with the Federal Reserve Banks, whereas the Federal Reserve notes rapidly get into general circulation. The deposits in Federal Reserve Banks get into general circulation only through the "deposit currency" issued by the member banks, which deposit currency is based on the reserves carried by member banks with Federal Reserve Banks. Since these "reserves" are in reality deposit accounts issued by the Federal Reserve Bank to member banks, it is clear that the deposit currency in general circulation is based on the "deposit reserves" issued by the Federal Reserve Banks. It should be noted, however, that the bulk of the reserves maintained by the member banks with the Federal Reserve Banks were originally established by the actual transfer of lawful money.

---

[1] On the average the reserves maintained by the member banks amount to about $10 for every $100 of demand deposits issued by them; and, since reserves of the member banks appear on the books of the Federal Reserve Bank as deposits against which a 35 per cent reserve of lawful money is required, it is clear that a minimum of $3.50 must be held against each $10 reserve deposit, which, in turn, represents $100 of demand deposits issued by the member banks.

Suppose, now, that the depositors in member banks are calling for cash, as in the Christmas season. At such times depositors have their deposits converted into cash. How can the member banks supply this cash? They can in turn demand that such surplus deposits (over and above their legal reserve requirements) as are carried with the Federal Reserve Banks be converted into Federal Reserve notes. The cash reserves of the Federal Reserve Banks will scarcely be affected at all by this demand for cash. Their deposit obligations have been converted into note obligations. But the reserves of the member banks *will* be affected. In this respect our banking system is not nearly so flexible as is the Canadian system, which permits every bank to do just what our Federal Reserve Banks alone can do; namely, to convert deposit obligations into note obligations. Our member banks cannot issue bank notes on a basis at all analogous to the basis on which they can create deposits. But the fact that our Federal Reserve Banks can do so gives a large measure of elasticity to the system. Our member banks may secure additional cash by rediscounting additional commercial paper and thus obtain Federal Reserve notes with which to pay their depositors. Thus member bank credit (individual deposits) is converted into Reserve bank credit (Federal Reserve notes). Indirectly, then, our system provides an elasticity similar to that of Canada.

Our Monetary System. Our circulating media thus consist chiefly of deposit currency and notes. Deposit currency composes the greater part of the *circulating media*, and Federal Reserve notes the major portion of the *money* in circulation. Back of deposit currency lie the cash reserves of the banks. In the case of banks which are members of the Federal Reserve system these reserves are deposited with the Federal

Reserve Banks, and back of these deposits in turn is the 35 per cent cash reserve in lawful money held by Federal Reserve Banks. Back of the Federal Reserve notes is the 40 per cent gold reserve held by the Federal Reserve Banks. Much of our money supply is therefore held as reserve against other money; for example, the gold that is held against Federal Reserve notes and United States notes. But more is held as a reserve against deposit currency. A dollar of lawful money held in reserve may become $30 in deposit currency. Therefore money proper cuts a much larger figure as a reserve against deposit currency than as a circulating medium. The reserve function of money is far more important than the circulating-medium function.

Gold is the base of our monetary system. It is our measure of value. Upon this base is built a superstructure of other forms of money and of deposit currency. The size of the superstructure depends upon two things: the quantity of gold and the quantity of money and deposit currency which the reserve efficiency of banks is capable of erecting upon each gold dollar without transgressing the limits of safety. Our circulating media can increase only as *gold* and the *reserve efficiency of banks* increase.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "In the latter part of 1862 the gold in the gold dollar became worth $1.15." How could this be possible? Explain precisely what is meant. How does Gresham's law apply to this situation?

2. If the market ratio in a bimetallic country was 17 to 1 and the mint ratio was 16 to 1, which metal would be driven out of monetary circulation?

3. "As a practical banker I cannot agree that bank deposits increase as a result of an increase in loans. In our bank we are always anxious to increase our deposits, since this enables us to increase our loans." Evaluate this statement.

4. A Minneapolis flour mill (Company X) sells flour to a merchant (Mr. Brown) in Winona. Suppose settlement is made with a bill of exchange. Write out a facsimile of such a draft. Suppose settlement is made with a bank draft. Write out the substance of such a draft.

5. "Banks are peculiar in this respect, that they are the only business institutions that boast of the volume of their debts." Indicate what is meant by this statement.

6. "Banks issue money and money substitutes. But the quantity that they can issue is limited." Explain.

7. "Banks may be said to be institutions which manufacture money. But this right ought to be reserved exclusively to the government." Do you agree?

8. "It is a great social waste to use such a valuable commodity as gold for monetary purposes. Government paper money would serve just as well." Discuss.

9. Over 9,000 banks are members of the Federal Reserve system. These banks control two-thirds of the banking resources of the country, but they constitute only one-third of the banks. Do you think that all the state banks should be compelled to become members of the Federal Reserve System?

CHAPTER XXI

PRICE FLUCTUATIONS

In this chapter we shall consider first the movements of prices historically, and, second, the effect of these price fluctuations on various classes in economic society. We leave to a subsequent chapter the consideration of the causes of these price movements.

Relative Prices, or Index Numbers. The changes in the price of a certain commodity over a period of time may be noted by comparing the actual price in a certain month or year with the actual prices in subsequent and preceding months or years. If, however, we are interested primarily not in the actual price, but in the *price changes* from month to month or from year to year, it will be preferable to use relative prices instead of actual prices. Relative prices show the percentage changes in prices from a certain base month or year. Thus, if the price in a certain year or month is 10 per cent higher than in the base year or month, the relative, or index number, would be 110, and the relative, or index number, for the base year would be 100. The device of relatives, or index numbers, makes it easier for the mind to comprehend quickly the degree to which the price in a certain period is above or below another period.

Relatives, or index numbers, perform another useful purpose: they make it possible for us to grasp readily the changes in the general level of prices of hundreds or even thousands of commodities. If one had before him the actual prices of several hundred commodities for a ten-year period, it would be absolutely impossible, even after extended study of the figures, to get any very definite idea of the extent to which prices generally had risen or fallen until these actual prices had been reduced to relatives, or index numbers. We shall now explain briefly how index numbers are constructed.

Table XIII gives the actual wholesale prices of five commodities from 1913 to 1923.

TABLE XIII. WHOLESALE PRICES OF FIVE COMMODITIES

	FLOUR (AVERAGE PRICE PER BARREL)	MILK (AVERAGE PRICE PER QUART)	BEEF (AVERAGE PRICE PER POUND)	MEN'S DRESS SHOES (AVERAGE PRICE PER PAIR)	ANTHRACITE COAL (AVERAGE PRICE PER TON)
1913	$4.42	$0.044	$0.130	$2.24	$5.31
1914	4.94	.042	.136	2.28	5.32
1915	6.50	.042	.129	2.48	5.33
1916	7.09	.045	.138	2.71	5.57
1917	11.23	.063	.167	3.28	5.94
1918	10.14	.078	.221	3.30	6.86
1919	11.98	.084	.233	4.32	8.27
1920	12.27	.085	.230	4.50	9.50
1921	7.97	.075	.163	3.21	10.53
1922	6.96	.073	.150	3.00	10.60
1923	6.19	.076	.158	3.15	10.88

If, now, we reduce these actual prices to relatives, or index numbers, we obtain a table like Table XIV. In this table we have selected the year 1913 as the base year with which all other years are compared.

TABLE XIV. INDEX NUMBERS, OR RELATIVES, OF WHOLESALE PRICES OF FIVE COMMODITIES
(1913 = 100)

	FLOUR	MILK	BEEF	SHOES	COAL
1913	100	100	100	100	100
1914	112	94	105	102	100
1915	147	95	100	111	100
1916	160	102	107	121	105
1917	254	143	129	147	112
1918	229	175	171	148	129
1919	271	191	180	193	156
1920	277	191	178	201	179
1921	180	169	126	143	198
1922	157	164	116	134	200
1923	140	172	122	141	205

From Table XIV we can readily tell at a glance the changes in price from year to year for each commodity; but how much had the prices of all five commodities risen on the *average* in 1917, for example, over 1913? One cannot easily tell this from the table. To answer this question it becomes necessary to construct an average index number from the five index numbers given. Now there are a great many different types of averages. Some give good results, and others give poor results.

The Making of Index Numbers. We shall consider only four types of averages: (1) simple arithmetical average, (2) weighted arithmetical average, (3) median average, and (4) simple geometrical average.

The simple arithmetical average is obtained by adding together the five index numbers (for the year 1917, for example) and dividing by 5. This type of average is defective in that it gives equal weight to each commodity, whereas they are as a matter of fact of unequal importance.

The weighted arithmetical average corrects this defect. It is calculated by multiplying each index number for any one year by its appropriate weight, adding the results, and dividing by the sum of the weights. By this method each of the commodities affects the average index number in proportion to its relative importance. If flour is relatively twice as important as milk, it should be weighted twice as heavily as milk. The Bureau of Labor Statistics weights each of the 550 commodities from which it constructs its index number of wholesale prices according to the total value of each commodity sold in the markets in 1923–1925. Each commodity thus affects the final index number in proportion to its relative importance.

If no weights are obtainable, it has been found that the median and simple geometrical averages are to be preferred to the simple arithmetical average. The median is calculated by finding the index number which occupies a central position with reference to all the rest. It is easy to calculate, since all one has to do is to arrange the various index numbers for any one year in order, from the lowest to the highest, and find the middle number. The simple geometrical average is the *n*th root of the product of the index numbers of *n* commodities. It is calculated by finding the logarithms of the various numbers for the year in question, summating these logarithms, dividing this sum by the number of logarithms summated, and then finding the antilog of this number.

These four types of averages are illustrated in Table XV. For the weighted arithmetical average the following weights were used: flour, 8; milk, 4; beef, 6; shoes, 1; coal, 5.

TABLE XV. AVERAGE RELATIVE PRICES OF FIVE COMMODITIES, 1913–1923

	SIMPLE ARITHMETICAL AVERAGE	WEIGHTED ARITHMETICAL AVERAGE	MEDIAN AVERAGE	SIMPLE GEOMETRICAL AVERAGE
1913	100	100	100	100
1914	103	105	102	102
1915	111	119	100	109
1916	119	124	107	117
1917	157	170	143	150
1918	170	181	171	166
1919	198	208	191	195
1920	205	214	191	202
1921	163	167	169	161
1922	154	156	157	152
1923	156	154	141	153

History of Prices. Fig. 25 shows the general movement of wholesale prices in England and the United States from 1790 to 1927. The curve for the United States is an average of about eighty commodities in the first part of the period and about four hundred commodities in the last part. The curve for 1790–1890 is constructed by the simple-arithmetical-average method, and the weighted arithmetical average is used for the period from 1890 to 1927. The British curve is constructed by means of the simple geometrical method from 1790 to 1850 and by the simple arithmetical method from 1850 to 1927. In spite of differences in commodities and method of construction the two curves show marked similarity, indicating that the forces affecting prices transcend national boundaries.

Let us center our attention on the American curve. Three large price upheavals are noticeable. The first was coincident with the War of 1812 and the Napoleonic wars in Europe, the second came with the Civil War, and the third with the great World War. Wars have been major causes of price upheavals.

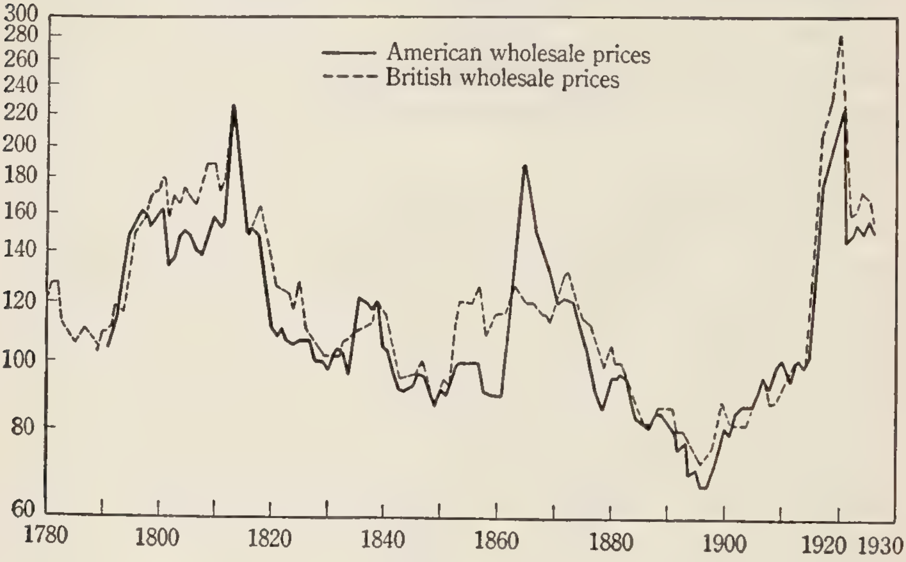

FIG. 25

Certain broad trends are noticeable. The general trend was upward from 1790 to 1814. Then there was a long downward trend from 1814 to 1849; then an upward swing from 1849 to 1865,[1] followed by a long downward movement which culminated in 1897. This was followed by an upward movement until 1920.

Superimposed on these long-run secular trends are the short-run fluctuations of prices which correspond to the movements of the business cycle. From 1895 to 1915 there were six price waves of this sort. Thus prices are constantly rising and falling, but each succeeding wave may be higher or lower than the preceding, depending on whether the long-run trend is in an upward or downward direction.

---

[1] The high point was reached in England in 1873.

Effect of Price Fluctuations. These price fluctuations constitute one of the gravest evils of modern economic life. The mere fact of high or low prices is itself of no consequence. Sometimes we hear people talk as though high prices were inherently an evil and low prices inherently good. To be sure, everyone wishes to buy at low prices and to sell at high prices; but since buying and selling are two sides of the same shield, it is obvious that this is impossible. If prices run uniformly on a high level, incomes from property and services may also be expected to run on a high level, and nobody is any worse off. Similarly, if prices are low, incomes from property and services will also be low, and nobody will be any better off. High and low prices in themselves make no difference. It is the *change* from high prices to low prices or vice versa that hurts. Rising prices benefit certain economic groups and injure others; falling prices injure some and benefit others.

It therefore becomes necessary in considering the economic effects of price fluctuations to deal with economic groups or classes. Numerous groupings might be made, but we shall concern ourselves with two pairs of opposing groups of major importance from the standpoint of price fluctuations. The first pair are borrowers and lenders. They are especially affected by the long-run trends of prices. The second pair are the entrepreneurial class and the wage-earning class, who are also affected by the long-run trends, but more especially by the short-run cycles. Since the borrowing class is chiefly composed of entrepreneurs, we have here really three groups — entrepreneurs, capitalists, and wage-earners. Entrepreneurs are chiefly affected as borrowers by the long-run trends of prices, and as employers and enterprisers by the short-run price cycles.

How do the long-run trends of prices affect borrowers and lenders? Consider first the borrowing class. Let us suppose that a certain farmer bought a piece of land in 1864 at a price of $20,000. He paid down $10,000 and borrowed the remainder on a fifteen-year loan. In 1879 the wholesale prices of farm products had dropped to a point less than half as high as the prices that prevailed in 1864. Under these circumstances the farmer's money income was rapidly falling. With a declining money income it became increasingly difficult for him to pay the interest on his mortgage and to meet the payment of the principal when it fell due in 1879. Nor could he renew his loan of $10,000 in 1879; for by that time the value of his land had fallen far below the 1864 level, and it would not have been safe for any investor to lend him three fourths or more of the value of his land. Not only was his income declining, but his original equity of $10,000 in the land was vanishing. By 1879 he had lost a large part if not all of what he had paid for the land. Falling prices were reducing both his income and his property.

Out of the decline in prices from 1865 to 1897 grew certain important social and political movements, such as the greenback, populist, and silver movements. Declining prices were robbing the borrowing class, and particularly the farmers, of income and property. No matter how hard the farmer worked and saved, the fall in prices defeated all his efforts. The dollar which he had borrowed had to be repaid in later years by a dollar which represented more products and therefore more labor and saving than the dollar he had borrowed. It was a "dishonest" dollar.

Consider, now, the effect of a rising price level on the borrowing class. Suppose a farmer bought a farm in 1900 for which he paid, let us say, $20,000. He advanced $10,000 of the purchase price and borrowed the rest on a ten-year mortgage. From 1900 to 1910 the general level of farm-product prices increased 47 per cent. The farmer's money income was rapidly rising. With his enlarged money income it proved much easier to pay off the loan of $10,000 than he had anticipated. He had borrowed dear dollars and was paying back cheap dollars. In the meantime the value of his land doubled. If our farmer was not disposed to save sufficient to pay off the loan, he would find no difficulty whatever in borrowing $10,000 on a farm now worth $40,000. Though he had not saved a cent of his yearly income, his equity in the land had increased from $10,000 to $30,000.

It is true that in the two periods which we have selected for illustration there were other factors, besides the fluctuations in the general price level, which affected the farmer's income and property. The development of the factory system and large-scale production, together with the progressive cheapening of transportation, were bringing manufactured commodities to the farmer's door at more and more favorable prices. Likewise the disappearance of free land and the great growth of population from 1900 to 1910 accounts in part for the fact that the prices of farm products rose faster than general prices. This affected favorably not only the purchasing power of the farmer's income but the value of all his property. This aspect of the problem will be taken up in greater detail in a later chapter.

Turn now to the lending, or investing, class. Suppose a man left his widow an estate of $60,000 in 1897 invested in bonds and mortgages. At 5 per cent interest this would yield her an income of $3000. From 1897 to 1914 the cost of living advanced about 40 per cent. The $3000 income would buy in 1914 only a little over two thirds of the goods and services which it would have bought in 1897. The rising price level had confiscated one third of the widow's real income. From 1914 to 1920 the cost of living more than doubled. The $3000 income could buy in 1920 only about one third of the goods and services that it could have bought in 1897. The rising price level had stolen two thirds of the widow's real income.

A striking illustration of the effect of a rising price level is to be found in Germany's experience during the decade from 1913 to 1923. Suppose a man had invested 4,000,000 marks (approximately $1,000,000) in high-grade bonds and mortgages. By August, 1923, prices had increased to such an extent that the purchasing power of 4,000,000 paper marks had been reduced to the equivalent of 6 pre-war marks. Thus a million-dollar fortune invested in the safest possible manner had dwindled in ten short years to a paltry dollar and a half. The rising price level had confiscated practically the entire fortune. Subsequently prices rose still more until in November, 1923, it took a trillion paper marks to equal a gold mark.

Now it must not be thought that this loss was nobody's gain. The physical property represented by these bonds and mortgages was probably still as valuable in real purchasing power as ever. The owners of this property who had borrowed on bonds and mortgages were the gainers. The rising price level *almost completely* wiped out their *entire* debt. Thus fluctuations in the price level take property from one class and hand it over to another class. Everyone recognizes the importance of legal protection of property rights, but few realize that fluctuations in the price level may reduce to an empty shell the property value ostensibly safeguarded by a bulwark of legal protection.

On the other hand, a falling price level favorably affects the investor. Every decrease in prices increases the purchasing power of his income and his property. His gain is the borrower's loss.

Consider now the short-run price cycles and their effect on the entrepreneurial and employing class on the one hand and the wage-earning class on the other. A period of rising prices is regularly a period of profit-making and prosperity for entrepreneurs, and a period of falling prices is a period of loss and depression. In the rising-price period all the entrepreneur's costs, with the single exception of raw materials, lag behind the prices of the products which he has to sell. But even on raw materials the entrepreneur often makes a profit in rising-price periods, since by the time they are worked up into finished products prices have risen. This argument would, however, not apply to orders taken in advance of manufacture. Wages and all the fixed charges, such as interest, rents, and salaries, lag behind general wholesale prices. With rising prices and costs lagging behind, the margin of profit is widened. Moreover, a rising-price period comes with an expanding market. Thus profits increase for two reasons: because costs lag behind selling prices and because the volume of sales is increased. Conversely, in falling prices the entrepreneur suffers a loss; for sales fall off and costs remain high, whereas selling prices rapidly decline.

The wage-earning class is affected in two ways by the short-run movements of prices: as to employment and as to wages.

Employment rises and falls with the cycle of prices. Since the market is improved and the margin of profits is greater in rising-price periods, entrepreneurs are eager to increase production and employ more labor. Similarly, in falling-price periods, when the market is restricted and profits are dwindling and often losses are suffered, the entrepreneur reduces his force and sometimes closes down completely. Thus employment rises and falls with the movement of prices.

The effect of rising and falling prices varies with different classes of labor. For our purposes it is sufficient to distinguish between common labor and skilled labor. Employment increases for all classes of labor in rising-price periods and decreases when the price level is falling. But the fluctuations are much more marked in the case of common labor than of skilled labor. When an employer is contemplating a reduction of his working force, he naturally decides to lay off his common labor first. He is loath to discharge a skilled man who may be hard to replace when business picks up. The skilled men constitute the vital part of the working personnel, and it is important to keep this part of the organization intact as far as possible. Thus the skilled man is kept on the job, and he is kept busy partly by giving him unskilled work to do even though he is paid at skilled rates. Common labor is fairly homogeneous; one man is about as good as another, and any one is easily replaced. Hence the employer does not hesitate to discharge the common laborers. They are the first to go, and with reviving business they are the first to be taken on again.

Many skilled laborers suffer greatly from seasonal unemployment, but common labor takes the brunt of the unemployment which goes with price fluctuations.

Price fluctuations affect not only the employment of labor but also the purchasing power of its wages. Here we are interested in retail prices and the cost of living and not in wholesale prices. Retail prices and the cost of living do not fluctuate nearly as violently as do wholesale prices. Not only is the rise and fall less extreme, but there is a considerable lag. Hence even though wages regularly lag behind wholesale prices and fluctuate less violently than these, it does not necessarily follow that wages fail to keep pace with the cost of living in rising-price periods or fall as rapidly in falling-price periods. Yet, on the whole, wages do fail to keep pace with the cost of living both on the upward and on the downward movement.

Here, again, the effect varies with different labor groups. So far as skilled labor is concerned, it is clear that wages fail to keep up with the cost of living in rising-price periods, but also fall less rapidly in declining-price periods. With unskilled labor the opposite is true: wages rise faster than the cost of living in the upward movement and fall faster in the downward movement.

Common laborers are therefore better off in the upward swing of the price cycle, both because of fuller employment and because of the higher purchasing power of wages. Likewise depression is an unmitigated evil for this group, both because of greater unemployment and because of reduced real wages. For the skilled laborers the case is not so clear. In the upward movement of the price cycle they suffer a distinct loss, since wages do not keep pace with the cost of living. On the other hand, they gain through fuller employment. In the downward movement they lose in diminished employment, but gain in so far as the cost of living falls more rapidly than wages.

But so far as industrial unrest is concerned, the effect of price fluctuations is clear. A rise in price gives rise to unrest for three reasons: first, the menacing rise in the cost of living, which, as we have seen, particularly affects the more aggressive and better organized laborers (the skilled); second, the growing volume of profits, which becomes generally known and leads to aggressive demands for higher wages; third, the increased demand for labor, which makes labor more independent and less wary of risking a struggle. On the other hand, falling prices also give rise to labor unrest. Wage cuts are resisted with widespread strikes, and unemployment gives rise to radical movements looking to a modification of the economic structure which makes such conditions possible.

Labor history is full of illustrations showing the influence of price fluctuations on industrial unrest. The first great labor movement of our country grew out of the enormous upheaval in prices from 1833 to 1837. National unions sprang up with the price upheaval of the Civil War and disappeared almost wholly with the panic of 1873. The membership of the American Federation of Labor fluctuates with the movement of prices, gaining with rising prices and falling off in declining-price periods.

Enough has been said to indicate that the problem of price fluctuations is one of the most important in the whole field of economics. We shall next seek to explain the causes of these price movements.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The real wealth of a country is not increased or diminished by fluctuations in prices. While these fluctuations are an injury to certain individuals, and a benefit to others, from a social point of view there is neither any gain nor any loss." Criticize.

2. "The price curve is the backbone of American history." — COMMONS. What is the significance of this statement?

3. "The depreciation of the monetary standard is a phenomenon of considerable social importance, and one whose effects must be regarded, on the whole, as beneficial. To begin with, its ordinary result is a rise in prices. Now a rise in prices is a useful stimulus to

production: it keeps the spirit of enterprise on the alert; it encourages a rise of wages; it acts as a tonic; it is a symptom of sound economic health." — GIDE. Do you agree with this statement?

4. "If all prices and incomes rose equally, no harm would be done to anyone." Is this statement true?

5. "Before 1896 the 'bloated bondholder' was gaining. Money lenders like Russell Sage rolled up wealth. They could not have done so after 1896. Even had they saved every penny of interest and compounded it, they would have had less actual purchasing power now than when they started. The newly rich today are not bondholders, but stockholders." — FISHER. Explain.

6. "Falling as well as rising prices cause discontent. For example, before 1896 the western farmer hated the eastern capitalist whose mortgages he found increasing in weight owing, he thought, to some manipulation of the market of money or produce or both." — FISHER. Evaluate this statement.

CHAPTER XXII

THE THEORY OF MONEY AND PRICES

In the last chapter we discussed price fluctuations and their effect on various economic groups; in this chapter we shall seek to explain the underlying causes of these price fluctuations.

The Value of Money. To begin with, we should note clearly the difference between fluctuations in value and fluctuations in price. It is conceivable that the price of a commodity might change very greatly, yet its value might remain unchanged. The value of a commodity is the quantity of *other commodities* for which it will exchange; the price of a commodity is the quantity of *money* for which it will exchange. Now if the prices of all commodities exactly doubled, no change would take place in the value of these commodities, but prices would rise 100 per cent. Money alone would change in value. A unit of money would now buy only half as much as before. The value of a unit of money, like the value of any other commodity, depends on the quantity of goods it will buy. Thus the value of money varies inversely with the general level of prices: when prices rise, the value of money declines correspondingly; when prices fall, the value of money rises.

Standard money may itself be made from a commodity upon which the government has placed its stamp through the process of coinage. Thus gold, which is our standard money, is itself a commodity. The value of standard money will always correspond to the value of the metal of which it is composed in a country that permits free (that is, unlimited) coinage of the standard money. If the value of the 232.2 grains of gold metal (and remember that by value we mean the quantity of other goods that it will buy) is greater than the value of the ten-dollar gold coin, gold as money will be melted up and converted into gold as a commodity, and vice versa. Thus in a country that is on the gold standard the *price* of gold never changes, but its *value* fluctuates inversely with prices. If the prices of all commodities (except gold) doubled, their value would be unaltered, but the value of the commodity gold would be reduced to half, while its *price* remained unchanged.

It is clear, then, that if *gold* is standard the money unit is tied to the commodity gold so far as its value is concerned. But it is conceivable that the standard money might consist of *gold coin* and yet *gold* might not be the standard of value at all. This could be true, however, only in case the coinage of gold were limited in quantity. In that case it is quite conceivable that a gold coin containing 232.2 grains of gold would have a value very much greater than 232.2 grains of gold. As a matter of fact, Austria at one time had as the standard money a silver coin worth considerably more than the silver it contained. In this case silver was just a commodity like any other commodity, fluctuating in price along with other commodities as the money unit appreciated or depreciated in value. The fact that the money happened to be made of silver had apparently no effect on its value. What, then, did determine its value?

The value of money depends in part on its abundance or scarcity as compared with the quantity of goods for which it is exchanged. Suppose the quantity of goods or commodities remains unchanged and the quantity of money doubles. Is it possible for the people of a nation to buy any more of the enjoyable things they desire simply because the quantity of money in circulation has doubled? Obviously not. It is clear that it is impossible to purchase any more goods and services than have been produced regardless of how much money may be in the country; and since there are no more goods to be purchased than before, it follows that more money will be exchanged for the same quantity of goods. This amounts to the same thing as saying that prices have gone up, or that the value of money has gone down.

Now it may have occurred to the reader to ask why a doubling of the quantity of money should cause any increase in the amount offered for goods. Might not the increase simply lie idle in banks, cash drawers, and private purses? Certainly, if the money were not offered in the market for goods, prices would not be affected. We must therefore distinguish sharply between the *actual quantity of money* in any given country at any one time and the *quantity of money offered in the market* during any given time. It is the latter that affects the value of money and the price level. When we speak of the "supply of money," we mean the quantity of money offered in the market in exchange for goods and services during any given time.

In the business world we say that we offer "money," even though we actually present checks or drafts which, as we have seen, are promises on the part of a bank to pay money on demand and are therefore as good as money. So when we speak of the quantity of money offered in the market in exchange for goods and services, we really mean the quantity of circulating media — deposit currency as well as money.

The Equation of Exchange. The quantity of circulating media offered in the market during a year in exchange for goods and services will depend on four things: (1) the quantity of money in circulation, (2) the number of times each unit of money in circulation is offered in the market during a year, (3) the quantity of deposit currency, and (4) the number of times the demand deposits have been drawn out and replaced. For these factors we shall use the following notation: M = the average quantity of money in circulation during the year (by "in circulation" we mean the money outside of banks and the reserves in the Treasury, and therefore available for the purchase of goods); V = the velocity of money, or the number of times each unit of money is offered for goods or services during a year; M' = the average quantity of demand deposits during the year; and V' = the velocity of demand deposits, that is, the number of times they are drawn out and renewed during the year.

Now, as we have seen in an earlier chapter, much of the money in circulation is based on reserves of gold or silver held in the vaults of the government and the banks. All the demand deposits are based on reserves of lawful money. Hence it is obvious that the quantities of M and M' are clearly related to the quantity of money held as reserves. Now if there is an addition to our stock of money either through an increased production or importation of gold or through an increased issue of government legal-tender paper money, this increase normally finds its way to the vaults of the banks and piles up as reserves. Large reserves result in low interest rates (unless interest rates are consciously controlled), which in turn stimulate borrowing at the bank. But an increase in bank loans, as we have seen, results in an expansion of demand deposits. Any large addition to bank reserves will therefore ultimately reflect itself in an increase in M and M'. Every addition to the monetary stock has in the long run resulted (at any rate until recently) in a larger and larger superstructure of circulating media built upon the foundation of these ever-expanding quantities of monetary reserves.[1] In fact, our deposit currency in this country has increased much more rapidly than our monetary stock. This increase in the ratio of our deposit currency to our total monetary stock was in part due to the fact that an increasing proportion of our monetary stock was being held as bank reserves, but more especially to a more intensive utilization of our cash reserves. Thus the superstructure of bank credit was growing even more rapidly than the reserves upon which it was based.

---

[1] If, however, the discount rate were regulated without regard to the accumulation of reserves (as appears to have been done by the Federal Reserve Board during the last five years), loans and deposits may not increase in spite of the piling up of reserves. The huge accumulation of gold in the Federal Reserve Banks during the last few years cannot properly be looked upon as an addition to our monetary stock. As Europe gradually returns to a gold basis our surplus gold will tend to flow out. The Federal Reserve Board has looked upon this huge gold surplus as a temporary phenomenon and has wisely refused to permit our discount policy to be controlled by this abnormal situation. Moreover, the vivid recollection of the evils suffered as a result of the recent boom and deflation has exerted a powerful restraint on bankers and business men.

All this amounts to nothing more than saying that loanable funds will not long remain idle. They may and often do lie idle temporarily in periods of depression, but the next swing of prosperity calls them into full use. The monetary and credit facilities of a country can in the long run regularly find full employment. Present purchasing power can always be utilized to advantage — if not in commercial loans, then in investment loans. It is inconceivable that additions to the monetary stock should merely result in piling up large bank reserves unable to find an outlet through loans in the form of bank deposits. Were this true, then we should have to conclude that loanable funds and present purchasing power are so abundant that there is no effective demand for them and hence the funds lie idle in the banks.

It is conceivable that if banking were conducted by a monopoly (or if the discount policy were closely controlled by a centralized body such as the Federal Reserve Board) loans and bank deposits would not increase in proportion to bank reserves. It might be argued that a banking monopoly could make bigger profits by restricting its loans and charging high interest rates. But such a curtailment of loans and bank credit would tend to produce a decline in the price level; such a policy would therefore be of doubtful benefit to the banking monopolist.

Government regulations of banking policy might in the interest of the public welfare refuse to permit bank credit to continue to be the football of accidental variations in the monetary stock and the bank reserves. It is quite possible, therefore, that we shall not find in the future such a close relationship between the total monetary stock and the quantity of circulating media as we have in the past; but in so far as these factors are permitted to work themselves out in an automatic, self-regulating fashion, we may expect every increase in the monetary stock to reflect itself in an increase in the circulating media.

V and V' depend largely upon (1) the density of population, (2) transportation facilities, and (3) the cyclical movement of business. If the population is becoming denser and the means for transportation more rapid and convenient, the velocity of the circulating media tends to increase. If business is good, money and deposits are rapidly being offered for goods; on the other hand, if business is depressed, money and deposits remain comparatively idle and are offered less rapidly in the market.

Prices rise, as we have seen, if an increased quantity of circulating media is offered in the market in exchange for a constant quantity of goods. But prices, or, conversely, the value of money, may also be affected by an increase in the volume of goods and services offered in the market. Suppose the volume of goods offered in exchange for money and its substitutes doubled while the quantity of circulating media offered in the market during the year remained constant, what would be the result? Clearly if double the quantity of goods and services were offered for the same quantity of circulating media, prices would fall to half the previous level unless the volume and velocity of the circulating media had changed.

Now the quantity of goods and services offered in the market depends fundamentally upon two things: (1) the total volume of goods produced during the year and (2) the number of times these goods have changed hands in raw, semifinished, or finished form. The former varies with the productive capacity of society, the yield of the harvests, and the condition of trade; the latter depends chiefly upon the degree of specialization as between regions, industries, and occupations. If ten firms constitute so many links in a chain of production from the raw material up to the finished product, and each sells his product to the firm constituting the next step in production, it is clear that the volume of *trade* is increased thereby, even though there were no increase in the volume of *production.* The volume of trade we shall designate by the letter T.

Now, as we have seen, prices tend to rise with every increase in the quantity of circulating media offered in the market, and to fall if the quantity offered decreases. On the other hand, prices tend to fall if there is an increase in the quantity of commodities (raw, semifinished, and finished) and services offered in the market, and to rise if there is a decrease in the offering of goods. Prices, therefore, vary directly with the quantity of circulating media offered, and inversely with the quantity of goods offered. We may express this relationship algebraically as follows:

$$\frac{MV + M'V'}{T} = P.$$

In this equation $MV + M'V'$ represents the quantity of circulating media offered, T represents the quantity of goods offered, and P represents the general level of prices. The relationship may also be expressed more briefly by the following equation:

$$\frac{C}{T} = P,$$

in which C represents the volume of circulating media offered in the market and T the volume of trade. It cannot be emphasized too strongly that the volume of circulating media, C, is affected as much by V and V' as by M and M'.

The effect of C and T on the price level can be illustrated historically in the two periods 1865 to 1897 and 1897 to 1920. In the former period T was increasing far more rapidly than C, and hence the general level of prices was falling; in the latter period C was increasing more rapidly than T, and hence the general level of prices was rising.

The Quantity Theory. We are now in a position to give a formal statement of a theory famous in economic literature — the quantity theory of money. John Stuart Mill, writing in 1848, put it as follows: "The value of money, other things being the same, varies inversely as its quantity, every increase of quantity lowering its value, and every diminution raising it, in a ratio exactly equivalent."

By "quantity of money" Mill means "the quantity of it which people are wanting to lay out; that is, all the money they have in their possession, except what they are hoarding, or at least keeping by them as a reserve for future contingencies." In other words, his terminology is equivalent to that which we have adopted above. "Supply of money" means the quantity of money offered in the market. "Money," says Mill, "acts upon prices in no other way than by being tendered in exchange for commodities. The demand which influences the price of commodities consists of the money offered for them."

He goes on to say that the "very same effect would be produced if we suppose the goods diminished, instead of the money increased; and the contrary effect if the goods were increased or the money diminished." And by goods being increased he means the goods offered in the market. Therefore he does not overlook the influence of trade on prices.

Nor does he overlook the influence of deposit currency. He says that "general prices at any moment depend much more upon the state of credit than upon the quantity of money. For credit, though it is not productive power, is purchasing power; and a person who, having credit, avails himself of it in the purchase of goods, creates as much demand for the goods, and tends quite as much to raise demand for the goods, as if he made an equal amount of purchases with ready money. . . . All this purchasing power, however, is operative upon prices only according to the proportion of it which is used." In other words, Mill measures the amount of credit by the quantity of it actually offered in the market.

As thus stated, the quantity theory of money is absolutely incontrovertible. No one can deny the validity of the equation of exchange which we have given above. It is a mathematical truism; for it must necessarily be true that the quantity of circulating media exchanged in the market must equal the quantity of goods exchanged multiplied by their prices. Thus $MV + M'V' = PT$. This equation merely points out a necessary relationship existing between six separate phenomena. A change in any one of them will necessarily result in a corresponding change in one or more of the other factors.[1] Let us suppose that P has increased. Then there must necessarily have been a corresponding change in some one or more of the other five factors. But which is cause and which is effect? Here is where the quarrel over the "quantity theory" begins.

Let us narrow down the range of the controversy. We can reduce our six factors to four. M and M' may be lumped together under the one head of M, or circulating media. The ratio between the quantities of these kinds of circulating media varies with the customs and conveniences of the people and depends upon whether money or checks are preferred as means of payment. If we lump M and M' together, it is obvious that we must likewise combine V and V' and consider the velocity of the circulating media as a whole. $MV + M'V'$ thus becomes MV. Our four separate and distinct factors are then quantity of circulating media (M), velocity of circulating media (V), volume of trade (T), and the general level of prices (P).

The quantity theorists argue that fluctuations in the quantity of circulating media (M) give rise to consequent movements in T, V, and P. From this point of view it follows that price fluctuations are ultimately to be explained by changes in the relative quantity of circulating media. On the other hand, the anti-quantity, or commodity, theorists argue that price fluctuations are altogether independent of changes in the quantity of the circulating media; that, as a matter of fact, the quantity of circulating media itself depends upon fluctuations in the price level, in the volume of trade, or in the velocity of the circulating media. Or, if we consider MV together, that is, the quantity of circulation media offered in the market (the *supply* of the circulating media as distinguished from its *quantity*), the anti-quantity theory holds that the supply of circulating media (MV) depends upon the general price level and the volume of trade.

---

[1] In so far as goods are bought and sold on book credit, the two sides of the equations are not necessarily equal. If we may assume that the total value of book credit does not fluctuate materially from year to year, this factor is not significant. In the last few years, however, it appears that there has been a considerable expansion of book credit as a result of the increase in *installment buying*. Sales on the installment plan amounted in 1925 to about \$6,179,000,000, on which \$1,638,000,000 cash payments were made, leaving \$4,541,000,000 to be paid on the installment plan. Over 60 per cent of the total value of these sales was for automobiles and trucks, nearly 30 per cent was for household goods, and only about 6 per cent was for clothing and jewelry. It should be noted that the dealer generally does not carry the customer himself, since, as a rule, he disposes of his installment paper to so-called finance companies or borrows more heavily at his bank. Thus installment purchasing is likely to give rise indirectly to an expansion of bank credit.

In considering these opposing theories it is necessary to separate the long-run secular trend of prices from the short-run cyclical price movements. Take the period from 1897 to 1915, when the long-run secular trend of prices was upward. How do these two opposing theories explain the upward trend of prices?

N.B. The anti-quantity theorists argue that this increase in the price level was not due to an increase in the quantity of circulating media as such. The change in the price level was due, they say, to the change in the value of the commodity gold. Gold was becoming more abundant as a result of newly discovered gold fields and because of improved processes in the extraction of gold. On the basis of our value theory it will be clear to the reader that as a result the marginal utility of gold as a commodity would decline. A unit of gold would therefore exchange for less of other things. Since our monetary system is based on the commodity gold (that is, on the quantity of gold contained in the gold dollar), it is clear that a gold dollar would exchange for less of other things. In other words, prices would rise because of the decline in the commodity value of the gold content of the dollar. In the meantime population and production were growing apace and the volume of trade was increasing. With higher prices and a larger volume of trade, more gold was being drawn into monetary use. Thus the monetary supply was increased not as a necessary consequence of the increased production of gold as a commodity but because the cheapening of gold raised prices, and this in turn made necessary a larger volume of circulating media.

N.B. On the other hand, the quantity theorists argue that the increased production of gold necessarily resulted in an increase in the quantity of circulating media. They hold that the value of gold as money is a distinct and separate thing from the value of gold as a commodity. As soon as the commodity gold became more plentiful and therefore cheaper, gold became worth more as money than as a commodity; hence gold necessarily flowed from the commodity market into the money market. The reserves of the banks were increased. These increased reserves readily gave rise to a larger volume of bank loans, since any available present purchasing power or loanable funds are soon fully utilized. Larger bank loans resulted in larger deposits, and thus the circulating media was inevitably increased. More and more circulating media being offered in the market for goods and services, prices rose. In this way the increase in gold production caused prices to rise.

Neither the quantity theorists nor their opponents lay any special stress on V, the velocity of circulation, in considering long-run trends. The quantity theorists would say that so far as the long-run trend is concerned, any changes in V depend mainly on changes in the customs and habits of the people, and that the dominant factors influencing the customs and habits of the people arise out of changes in their economic environment, such as increasing concentration in population, improvements in the means of transportation, and the development of banking institutions. With this their opponents would in the main agree, but they would be inclined to add that the velocity of circulation may be stimulated by an upward trend in prices and an increasing volume of trade, particularly if the quantity of money needed to move the goods at the prevailing prices were not readily forthcoming.

MV = P

Quantity theory Increase T results in lower P.

anti-Quantity Increase T may result in increase M.

T

With regard to the volume of trade the two theories take a radically different point of view. The quantity theorists hold that an increase in the volume of trade has the same effect as a decrease in the supply of circulating media. The term "supply of circulating media" must always be taken in a relative sense. Thus, if the increase in the volume of trade equals the increase in the volume of circulating media, no change in the price level will follow. The quantity theorists would not deny, however, that the volume of trade might itself be stimulated by an upward trend in prices, but any such increase in the volume of trade would to that extent check the upward movement of prices.

The anti-quantity theorists would not admit that an increase in the volume of trade necessarily tends to lower prices. They would admit that if the gold supply remains constant and the supply of all other sorts of finished goods and services is increased, the exchange ratio will become more favorable to gold, and prices will fall. But if the volume of trade is increased merely because of an increase in the intermediate steps in the process of production, such increase in the volume of trade would have no effect on the price level, according to this theory. It will, to be sure, call out a larger volume of circulating media, but that would have no effect on the general price level.

It is clear, then, that the quarrel centers on *M* and *P*. One side argues that fluctuations in the circulating media result in price fluctuations; the other side argues that price fluctuations cause fluctuations in the quantity of circulating media.

Both sides agree that there is a correlation between prices and circulating media. But what is the causal relationship? Which is cause and which is effect?

When Austria issued a silver coin in limited quantity as the standard of value, the result was that the coin became worth more in exchange than the silver contained in it. India furnishes another example of the same phenomenon. These facts are readily explained on the basis of the quantity theory.

On the basis of the anti-quantity, or commodity, theory we should expect that standard metal money would always be worth exactly the same as the quantity of metal of which it was composed. Here, then, are historical facts which the quantity theory can explain and which are not explicable on the basis of the commodity theory.

The case of irredeemable paper money is not so clear. From 1862 to 1879 we were on a greenback-dollar basis. What determined the value of that dollar? The quantity theory answers that the value of the greenback dollar depended solely upon the quantity put into circulation except in so far as this factor was counteracted by other factors, such as a change in the volume of trade etc. That a close relationship held between the quantity of greenbacks in circulation and the general price level can be proved statistically. But the opponents of the quantity theory contend that this statistical correlation proves nothing, that the real cause of the depreciation of the paper dollar (and hence of the rise in prices) was not the quantity in circulation but the prospects of the redemption of greenbacks in gold. They hold, of course, that the more the greenbacks increased, the darker became the prospects of gold redemption and the more rapid the rise in prices. It was not, they say, the quantity of greenbacks per se, but the effect of an ever-increasing quantity on the probability of early redemption in gold that caused prices to rise.

Now it is scarcely likely that the faith of people in the redemption of the greenback could fluctuate with the smoothness and steadiness exhibited by the curve of general prices during this period. The confidence of the people in the redeemability of the greenback was certainly affected by many factors besides the quantity in circulation. The stability of the government and the success or failure of the Union armies were important factors. These factors had little effect on the price level, but they profoundly influenced the confidence in the redeemability of the greenback as evidenced by the fluctuation in the exchange ratio between greenbacks and gold.

Gold fluctuated far more in value than greenbacks; it appreciated in value when the prospects were that gold would again become the standard money, and vice versa. These facts tend to support the quantity theory.

Nor does it seem reasonable to suppose that the depreciation of currencies in central and eastern Europe in 1922–1923 could be explained on the probability of gold redemption. It would seem more correct to say that the probability of gold redemption was absolutely nil. On the other hand, a fairly close relationship did exist between the volume of the circulating media and the general price levels in the various countries.

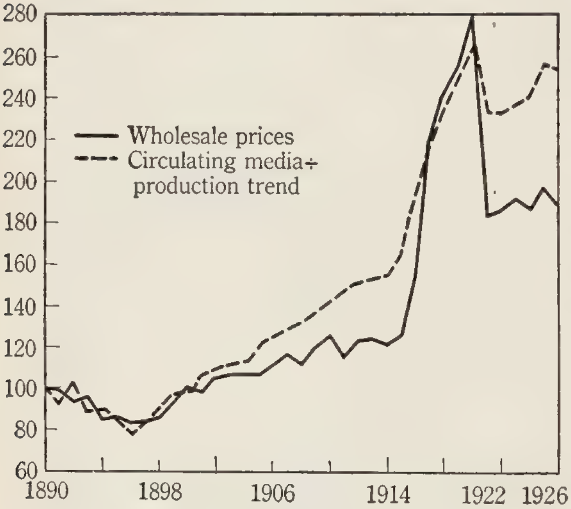

FIG. 26

Fig. 26 shows the statistical relation between the fluctuations in the circulating media[1] (corrected for the constantly increasing production trend) and wholesale prices in the United States. The circulating-media index represents a weighted average of money in circulation and deposit currency. This index has been divided by an index of the *trend* of production. If we designate the weighted circulating-media index as M, the production trend as G, the wholesale price index as P, and substitute G for T in our abbreviated equation of exchange, we have $\frac{M}{G} = P$.[1] It will be seen from the chart that the price index, P, lags behind the corrected circulating-media index, $\frac{M}{G}$, particularly after 1900. This is probably due chiefly to the fact that the volume of trade, T, increased more rapidly than production, G, owing to increased specialization, both geographical and technical. This necessarily implies an increased exchange of goods produced and hence an increase in the volume of trade exceeding the increase in production. Thus we should expect $\frac{M}{G}$ to outrun P. However, during the World War inflation period, P outran $\frac{M}{G}$. This may probably be explained in part on the ground that during a period of inflation the velocity of money and bank deposits increases greatly, and this increase in velocity would tend to raise prices more than would be indicated by the increase in $\frac{M}{G}$. It may be explained in part by the fact that wholesale prices do not constitute a perfect measure of the general price level. A general price level should also include wage rates, security prices, retail prices, rents, transportation rates, and the like. Wholesale prices rose higher in the period of great inflation than these other items in the general price level. Thus the curve given in the chart exaggerates the actual rise in the general price level. By 1921, however, the price level had again fallen behind the circulating-media level, thus approaching again the prewar relationship.

---

[1] The "circulating media" is a weighted average of money in circulation and demand deposits. The figures for demand deposits are given by Holbrook Working (*Quarterly Journal of Economics*, February, 1923, and *Review of Economic Statistics*, July, 1926) and Allyn A. Young (*Review of Economic Statistics*, July, 1927). The production trend was estimated from the indexes prepared by W. W. Stewart (*American Economic Review*, March, 1921), E. E. Day (*Review of Economic Statistics*, 1921), and W. F. Maxwell (*Review of Economic Statistics*, 1927).

KP. So far as the short-run cyclical fluctuations are concerned the argument of the anti-quantity, or commodity, theorists is somewhat shifted. The long-run swings of prices, it is said, can be explained by the appreciation or depreciation of the commodity gold. The value of the commodity gold establishes a normal price level. But the short-run ups and downs of prices are due to other factors. They cannot be explained in terms of the rise and fall in the value of gold (in gold-standard countries), nor can they be explained, from the standpoint of the commodity theory, in terms of fluctuations in the circulating media. These cyclical movements of prices are the result, it is said, of the forces that produce the short-run fluctuations known as the business cycle.

---

[1] According to Carl Snyder the *cyclical* fluctuations of V and T tend to offset each other. Hence we should expect that the volume of circulating media divided by the production *trend* would correspond closely with the movements of the price level.

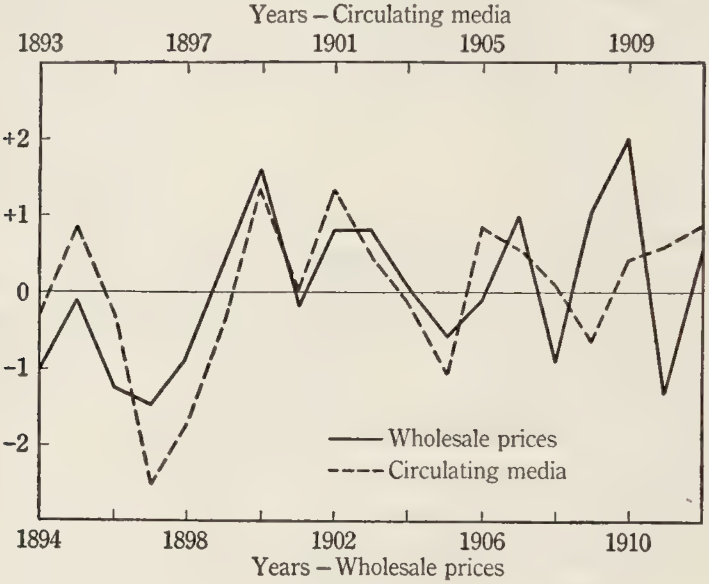

FIG. 27

On the other hand, some quantity theorists argue that the short-run movements of circulating media explain the cyclical movements of prices.[1] They hold that these price cycles in turn produce or constitute the business cycle. The argument is therefore just the reverse of that of their opponents.

In the accompanying chart of prices and circulating media (Fig. 27) the long-run trends have been eliminated. We have, then, the short-run, or cyclical, movements of the two series.

---

[1] The explanation is, however, not complete, since the movements of the circulating media need to be explained. For a complete explanation see R. G. Hawtrey's theory of the business cycle, as explained on pages 387–388 of this book.

The chart shows, except for the last few years, a close correspondence between the two series; but since the curve for circulating media is lagged one year on the chart, it is clear that changes in the circulating media actually precede changes in the price curve by one year. When an increase or a decrease takes place in circulating media, prices do not respond instantly.

The fact that prices lag behind the fluctuations of the circulating media would seem to lend support to the quantity theorists. The opponents of the quantity theory contend, however, that the lag of prices proves nothing. The movement of loans and deposits is, indeed, a necessary condition but not a cause of the movement of prices. Thus, if it is anticipated that business profits will rise next year because of new inventions, large crops, the discovery of new resources, the opening up of new markets, or the increased demand for goods consequent upon the outbreak of a war, optimism is engendered, business men will borrow more heavily from the banks, and the volume of deposit currency will rise. Thus the increase in loans and deposits is the result of forces affecting the anticipated rate of profit. These "forces" are the real causes of the increase in prices; the increase in loans and deposits are merely the mechanism through which these forces operate.

The Income Theory. Albert Aftalion, a French economist, has recently advanced an income theory of prices.[1] Price fluctuations, he believes, depend upon the respective movements of the *money* income and of the *real* income, or, in other words, income measured in terms of commodities and services. If the *money* income of a given community increases while the *real* income remains the same, prices will rise; if the money income remains stationary while the real income is increased, prices will fall. This relationship may be stated by means of the equation $I = PR$, in which I is the money income, P the price level, and R the real income in terms of goods and services.

---

[1] *Revue d'economie politique,* May–June, 1925. Aftalion here carries forward a suggestion made by F. von Wieser in the *Schriften des Vereins für Sozialpolitik* in 1909.

This theory has the advantage that it runs in terms of the modern theory of value with its demand-and-supply schedules. If the money income of a given economic society is increased, the demand schedules for goods rise and prices are forced up.

The demand schedules for goods are in reality money-supply schedules; that is, offers of various amounts of money for different quantities of goods. As the money income increases, the amounts of money offered against given quantities of goods increase. This statement of the case is in no way contradictory to the quantity theory as we have explained it, but it has the advantage of stating the effect of increases or decreases in money income in terms of the supply-and-demand schedules for the numerous individual goods that are bought and sold in the market. Moreover, it views the problem less mechanically than the quantity theory does. According to the quantity theory we might expect that a certain increase in money income would increase the general level of prices regardless of which sections of the community receive the increased income. But according to the income theory the increase in prices will be different, depending upon who the individuals are whose money income is increased. What classes of the people have received larger money incomes? rich or poor? savers or spenders? demanders of luxuries or demanders of necessities? The increase of incomes does not affect the price mechanically, or automatically, but through the desires of those whose incomes are raised.

Yet even here the income theory is not at variance with the quantity theory, properly understood. If the quantity theory is stated mechanically, and the assumption is made that with a given increase in money income everything else remains the same, it follows, of course, that the quantity theory, so stated, leads to wrong conclusions. In point of fact everything else does not remain the same. If the money income of the rich increases while the money income of the poorer class remains the same, shifts in demand for the different products will take place, and in consequence the relative production of different goods and services will be altered; if, on the other hand, the money income of the poorer class is increased while the money income of the rich remains the same, the shifts in demand will be of quite a different nature, and the subsequent shifts in production will be in favor of goods purchased by the poorer class. Now a change in production affects, as we have seen, the price level just as surely as does a change in money income. Since the change in production is not the same when the income of the poorer class is increased as it is when the income of the rich is increased, it follows that a given increase in money income will not affect the general price level to the same extent in both cases. Now one of the main objections to the equation of exchange is that it is likely to mislead us into viewing the problem as though we were dealing with a static society. We are likely to assume that "other things remain the same," when in point of fact they do not. It is quite possible that the income formula gives a more dynamic approach to the problem of price fluctuations. Possibly, also, the income formula may help us to apply the equation of exchange with better insight into the relationships of cause and effect between the various factors which it takes into consideration.

Stabilization of the Value of Money. Professor Irving Fisher has advocated a far-reaching monetary reform to stabilize the value of money. In place of a standard gold dollar of *constant weight* but *fluctuating value,* such as we now have, he proposes a dollar of *constant value* and *fluctuating weight.* This dollar would not be coined, but would circulate in the form of gold certificates. These certificates would represent a varying amount of gold, depending upon the prevailing content of the standard gold dollar.

A properly authorized governmental bureau would be required to change every two months the weight of the standard dollar with a view to maintaining a uniform price level. Suppose the system were adopted this year. We would start off with a gold dollar weighing 23.22 grains of fine gold. In two months prices have risen, let us say, 1 per cent. The bureau would then increase the weight of the standard dollar 1 per cent. This would tend to check the rise in prices. Every purchaser would now be exchanging a dollar containing 23.45 grains of gold. Obviously a larger gold dollar would buy more goods. If more goods can be purchased for a dollar it is obvious that prices have fallen. At the end of another two months it is conceivable that prices would have fallen to a point half of 1 per cent below the price level prevailing in the initial, or base, period. The bureau would then reduce the weight of the dollar by half of 1 per cent, making its weight 23.33 grains. Thus at periodic intervals the gold content of the standard dollar would be increased or decreased with a view to stabilizing prices. The index number of prices would thus fluctuate 1 or 2 per cent above or below the initial price level, unless violent fluctuations precluded adjustment.

To avoid speculation in gold at the expense of the government, a small fee would have to be charged all persons who exchanged *gold* at the Treasury for gold certificates; otherwise a person might deposit 23.22 grains of gold with the Treasury just the day before a probable increase in the weight of the standard dollar would be announced. He would receive, in exchange for the 23.22 grains of gold, a gold certificate entitling the bearer to the gold content of one standard dollar. On the next day the gold content is increased say 1 per cent. He would then be able to get 23.45 grains of gold for his gold certificate. As a speculator in gold he would thus have made a profit overnight of 1 per cent. To prevent this, no single change in the weight of the dollar should exceed the fee charged gold-depositors.

Fisher's proposal, it will be noted, does not depart from the gold standard of value. The value of standard money would still, as now, depend upon the value of its gold content. If gold was being mined or imported in large quantities, his plan would serve to bottle up more and more gold in the government Treasury. Instead of the new gold's becoming the basis for an increase in circulating media, more gold would be put back of every gold certificate in circulation. Fisher's scheme would therefore regulate the *quantity of the circulating media*, adjusting it to the demands of trade in such a way that the price level would be stabilized. At the same time the plan would have the merit of retaining gold as the ultimate standard of value. The public, long accustomed to the gold standard, would certainly have more confidence in a monetary system based on gold than in one based entirely on the scarcity value of a limited circulation. Whatever may be true of the future, a gold-standard monetary unit is still safer and more acceptable than paper money, the value of which rests entirely on its legal-tender quality and on the limited quantity issued.

An objection frequently raised against Fisher's plan is that if one country adopted it alone, it would experience difficulty in foreign trade because of the fact that while its money was stabilized the monetary units of other nations would continue to fluctuate. Thus the exchange rates would constantly be changing. This objection is probably not very serious. During the World War and post-war periods exchange rates have fluctuated greatly because of the fact that most European countries were on a paper basis while we were on a gold basis. Yet this fact has not proved a serious obstacle to trade relations. Moreover, before the war exchange rates were constantly fluctuating above and below par. The worst that can be said is that if we adopted Fisher's plan and Europe did not, the fluctuations in the rates of exchange would be somewhat greater than they now are between gold-standard countries.

Stabilization of the Value of Gold. Another proposal is to stabilize the value of gold itself. This could be accomplished in one of two ways: (1) by regulating the production of gold and (2) by controlling the demand for gold.

Professor R. A. Lehfeldt has made the proposal that we undertake to regulate the supply of gold coming from the mines in order to stabilize its value. This might be accomplished through the establishment of an international gold monopoly. If gold was being mined too freely and its value was therefore falling (that is, the general price level was rising), the output of gold should be curtailed. If, on the other hand, gold was appreciating in value (the general price level falling), the production of gold should be stimulated even though this required the mining of low-grade ore at high cost. In this way the value of gold would be stabilized.

Gustav Cassel has estimated that the monetary stock of gold increased from 1850 to 1910 at the average rate of 3 per cent per annum. The volume of trade was apparently increasing at a somewhat greater rate, but in view of the increasing banking efficiency and the greater superstructure of deposit currency, based on a given quantity of gold, a 3 per cent rate of increase was apparently sufficient to prevent a downward trend in prices. At present the annual gold production is approximately $400,000,000 per year, half of which goes into new monetary uses. The total monetary stock of gold in the world (about $10,000,000,000) is therefore increasing currently at a rate of only about 2 per cent per annum. The prospects are that the output of gold will decline still further, in ten or fifteen years, with the progressive depletion of the important gold fields of South Africa. It appears, then, that even though it were feasible to organize an international gold trust, it might be difficult to secure a sufficient output of gold to prevent a fall in the general level of prices.

The second proposal is to stabilize the value of gold by regulating the demand for it. Professor E. W. Kemmerer has suggested that it might be feasible to restrict the flow of gold into nonmonetary uses by taxes and other governmental measures. Moreover, the monetary demand for gold might be curtailed, as advocated by Cassel, in the following ways: (1) keeping gold out of general circulation and using it exclusively for reserve purposes, (2) reducing the customary gold reserve ratios, and (3) centralizing the gold reserves of the world in one or two centers, such as New York and London. Moreover, a great economy in the use of gold for monetary purposes could be effected if the continental European nations could be educated to use checks (demand deposits) more freely. In this manner the threatening scarcity of gold may be offset by a curtailment in the monetary demand for gold, thus preventing a progressive fall in the general level of prices.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It has sometimes been assumed that the quantity theory is obviously fallacious since we know from history that the quantity of money has frequently increased and yet there has been no increase in prices, and, conversely, the price level has at times changed without any corresponding change in the quantity of money. Similarly, it frequently happens that the supply of a certain commodity has increased without any corresponding change in its value. Is it safe to conclude from this fact that the law of supply and demand has no validity? What possible factors have not been taken account of in both the cases referred to?

2. "But if we divest ourselves of all prejudices, and minimize the effects of local and temporary conditions by making long-range studies, in numerous countries, of price levels and money quantities, our search brings us to a conclusion from which there is no escape. We discover that there have been no large, sustained movements in prices without corresponding changes in the relation of the volume of money to the volume of goods." — FOSTER and CATCHINGS. Illustrate from American economic history.

3. "If fresh means of payment are created — whether in the form of legal tender currency or in the form of banking means of payment — additional purchasing power is created. The result must be a rise in prices. We cannot say exactly how great this rise will be." — CASSEL. Why not?

4. "Between May of 1920 and December of 1924 the volume of monetary gold in the United States increased 73 per cent and the net demand deposits of reporting banks increased 14 per cent. Commodity prices at wholesale, however, dropped 37 per cent during this period. During the same period changes occurred in British money,

credit and prices which the quantity theory cannot possibly explain." — ANDERSON. Do you agree? Why, or why not?

5. "Obviously there is a great shortage of money in the United States, since the property value of the country is about twenty times as great as the circulating media. Every dollar of wealth ought to be matched by a dollar of circulating media." Criticize.

6. "The greenbacks depreciated in value not because of the large quantity issued, but because of the loss of public confidence in the ability of the government to redeem the greenbacks." Evaluate this argument.

7. "Before the war the per capita quantity of money in circulation in France was about five times greater than in England, yet the per capita volume of business was probably greater in England. On the basis of the quantity theory we should expect prices to be about five times higher in France than in England. Obviously there is something wrong with the quantity theory." Criticize.

CHAPTER XXIII

BUSINESS CYCLES

Industrial Fluctuations. The wholesale-price chart which appears on page 340 shows three types of price fluctuations: first, it reveals three periods of extraordinary price inflation; second, it shows unmistakably five long-run trends in prices — three upward trends culminating in 1814, 1865, and 1920, and two downward trends culminating in 1849 and in 1896; third, it shows numerous short-run fluctuations, which indicate the ups and downs of business prosperity and depression. It is with these short-run fluctuations that we are concerned in this chapter.

The term "business cycles" is of recent origin. Formerly economists and business men centered their attention on crises, panics, and depressions; they did not think of industry, commerce, finance, and business in terms of *cycles*. The increasing application of statistical research in the field of economics has revealed unmistakably that business does not normally run along on an even plane of prosperity only to be thrown suddenly from time to time into deep gulches of depression. On the contrary we find that business and industry are constantly rising and falling in more or less rhythmic waves, or cycles. Whether we look at the physical volume of production and sales or at the prices of economic goods, we observe this same cyclical rise and decline. These wavelike movements, or cycles, appear in every field of industry, commerce, and finance. They may be seen in the production of such basic materials as pig iron or coal, in the volume of construction work, in the nations' exports and imports, in the employment of wage-earners in factories, in the volume of manufacturing production, in the volume of freight carried by the railroads, in the prices of commodities sold at wholesale and retail, and in the changes in wage rates. On the stock market they appear in the prices of stocks and bonds and in the volume of sales of securities. In the field of finance they are evident in the movements of loans, deposits, reserves, discount rates, commercial-paper rates, call-loan rates, bank clearings, and liabilities of business failures. In fact the whole range of business activities runs in wavelike movements or cycles.

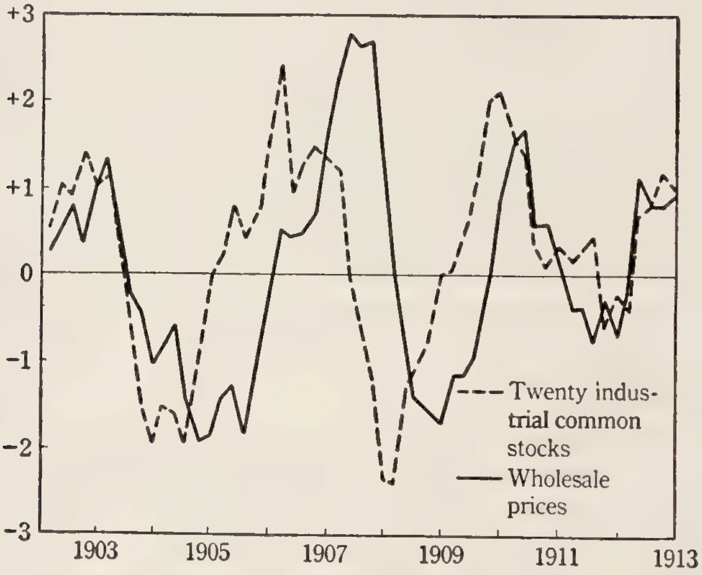

FIG. 28

Now it must not be thought that the mere fact that these different phenomena of economic life show up-and-down fluctuations constitutes in itself any proof that there is such a thing as a business cycle. If these fluctuations were merely sporadic oscillations back and forth without relation to each other, they would not constitute a business cycle. In reality, however, we find that the wave-lengths of these various movements are substantially equal and that there is close correlation between these fluctuations. This at once establishes a strong presumption that these various movements are all the result of a common set of forces which operate to produce what is known as the business cycle.

But while these fluctuations have a common wave-length they are not all synchronous. Thus the prices of the shares of stock on the speculative exchanges usually forecast by several months the conditions of business represented by such indexes as wholesale commodity prices, the production of pig iron, railroad traffic, and employment in factories.

Various Phases of the Business Cycle. The business cycle has several phases, which may be denominated as follows: (1) the prosperity phase, (2) the crisis phase, (3) the depression phase, and (4) the recovery phase. We shall now describe in some detail the chief characteristics of each of these phases. In the prosperity phase of the cycle we observe the phenomena of rising prices. The selling prices of the finished products of manufacturers and the stocks of merchants rise more rapidly than costs. Wage rates do not start upward until from six to nine months after wholesale prices. Interest rates on current loans lag from three to six months behind the movement of prices. All the overhead expenses of the business remain relatively fixed. Rents are generally fixed by long-term contracts. The interest charge on outstanding bonds remains fixed until the bonds mature, generally a period of ten, twenty, or thirty years. Salaries change very slowly in spite of a general upward movement of prices. Freight rates and the rates on power and light, regulated as they are by public authorities, move slowly and tardily during the major portion of the upward movement. Although the prices of raw materials rise fully as rapidly and as high as the wholesale prices of the finished products, yet the producer who manufactures for stock enjoys a profitable spread between the prices paid for raw materials and the prices subsequently received for the finished product. Manufacturers, retailers, and wholesalers find their inventories rising day by day in value. So the spread between costs and selling prices is widened. Profits rise not only because of this favorable price margin but also because of the increase in the volume of production and sales. New plants and additional equipment are installed. Retailers, wholesalers, and manufacturers, encouraged by rising prices and growing markets, increase their stocks of goods. Expansion is the order of the day. Rising prices, lagging costs, a wide margin of profit, and large inventories are the characteristic features of the prosperity phase of the cycle.

But the time is soon reached when factories are running at full capacity and no further increase in physical production is possible. The competitive bidding for goods and labor, however, continues, made possible by further borrowing at the banks and the consequent further increase in the volume of circulating media. In short, the purchasing power of the community measured in terms of money units continues to rise in spite of the relative limitation in the volume of goods available for purchase. Two results follow from this: prices continue to rise still further and an increasing number of orders remain unfilled.

During this phase of the cycle the demand schedules of all sorts of buyers are rising. Buyers, especially in dealers' markets, are willing to pay more and more money for the *same* quantities. The demand schedules are moving upward, since the banks are furnishing, through the increase in loans and deposit credits, more and more purchasing power measured in terms of money units.

There is, however, a limit to the possible expansion of the community's purchasing power measured in terms of money units. The reserves of the banks finally approach the danger point and ultimately the limits imposed by law. Fresh loans can no longer be made, and the expansion of bank credit comes to an end. A check is thus placed upon the competitive bidding for goods and services. The upward movement of prices is checked. Business men who hold large stocks of goods become nervous, and finally, fearing a fall in prices, throw their stocks on the market. A general selling scramble ensues, and prices drop.

It may be that the upward movement is checked before the final limitations imposed by the reserves of the banking system as a whole are reached. For example, in 1923, when the member banks had reached the limit of their own capacity to extend loans they were reluctant to extend their credit further by rediscounting at the Federal Reserve Banks. Moreover, the officers of the Federal Reserve Banks and the Federal Reserve Board were reluctant to encourage heavy rediscounting, since they feared a further period of huge inflation such as we experienced in 1919 and 1920.

Other factors may impose a check on the expansion of bank credit. A bad harvest may make business men fearful of the future and cause them to hesitate to borrow further. Certain lines of business may be overdeveloped, with a consequent fall in profits. This affects the stability of interrelated industries and causes business men to lose confidence in the future. A war will have its effect, more or less extensive, depending upon its scope. All these factors affect the course of the cycle and produce marked differences between various cycles.

Gradually business swings into the crisis phase of the cycle. The demand for goods ceases to rise and begins to decline. Prices first rise less rapidly and then begin to drop. This gives the cost items, which got a slow start in the upward movement, a chance to catch up with selling prices. Wage rates continue to rise after the price movement has begun to slow down. Discount rates on commercial paper rise sharply. Extensions of plant and new equipment installed in the prosperity phase of the cycle and purchased at high prices now become a heavy load for the company to carry, for the fixed expenses entailed by these extensions must be met each year in spite of the new turn of affairs. Salaries, rents, and other overhead expenses have finally been pushed up late in the prosperity phase of the cycle. When the market begins to recede and prices turn downward, these heavy charges on business remain as a millstone round its neck. With costs remaining high and prices falling, the profit margin is converted into a loss margin. The great problem for business men becomes not so much to make profits, but rather to prevent too heavy losses. In fact, the first task confronting business men when they become aware that the crisis has really come is to maintain their solvency. Not having anticipated the crisis, they find themselves tied up with numerous contracts and commitments, with loans and obligations entered into under the expectation of sales and prices which cannot now be realized. They must pay their debts or risk being pushed into bankruptcy. Hence accumulated stocks are thrown on the market at whatever prices they will fetch, and this intensifies the fall in prices. The receipts are not nearly sufficient to meet obligations. Accordingly there ensues a pressure at the banks for new loans and extensions of time on past loans as they mature. Each business man finds it impossible to collect from his creditors or to pay his own obligations. Credits thus become "frozen"; debts are not paid when due. Gradually, through rigid economy and saving, the credits begin to thaw out. This is the process of "liquidation." The final phase of the liquidation process consists in forcing into bankruptcy those who are totally unable to liquidate, thus "thawing out" by legal pressure the hardest-frozen credits. This last phase of the liquidation process becomes most widespread after the crisis is passed and industry has passed into full depression.

When business men become thoroughly aware that their expectations are not going to be realized because of the decline in the market and the fall in prices, the first job that confronts them is to maintain their solvency. This is the outstanding characteristic of the crisis phase of the cycle. It is chiefly a financial problem. As business passes gradually into a period of more or less prolonged depression the problem of production and sales comes to the foreground. The manufacturer is confronted with the prospect of a falling market; each new quotation brings news of a further fall in prices. He is confronted with high costs on every hand: high wages, large fixed interest charges on past loans, high rents, high salaries, high taxes, high freight rates and public-utility charges. If he continues to operate, he may find his costs higher than his selling prices. At first thought it might seem that the proper solution is to cut the Gordian knot by closing up his plant entirely. But unfortunately for him, although his gross income would be nil, his costs would in large measure still continue should he resort to this drastic action. He could, to be sure, eliminate his wage and salary bill, but he hesitates to discharge skilled, trusted, and responsible men; he must keep the essential part of the organization together. In no way can he get out from under the fixed interest charges, the rentals, and the taxes. Moreover, he has a considerable stock of raw materials on hand which are constantly falling in value. This stock at least must be cleaned out before it is feasible to close up shop. When the old inventories have been worked off, some plants close down completely for a few weeks or longer; others keep on running at still further reduced capacity.

Every economy is resorted to in the utilization of office supplies, plant equipment, power, light, and materials, and every wage-earner or clerical worker who is not highly essential is discharged. Wages and salaries are cut down, though (except for unskilled labor) not in proportion to the fall in selling prices.

Merchants are faced with a situation fairly similar to that which confronts manufacturers. When the drop in prices comes, they are faced with the problem of what to do with their huge stocks bought at high prices. At first they endeavor to work off as much of their stocks as possible at the old prices. Finally certain merchants, desperately hard pressed for cash, cut prices considerably in order to liquidate a material part of their stocks. Others, seeing what is coming, cut prices in order to maintain the physical volume of sales. Eventually new stocks come on the market at lower prices, and all retailers are forced to reduce prices. In the face of this declining market, merchants, like manufacturers, endeavor in every possible way to push down costs.

As business proceeds into the depression phase of the cycle, bank loans and deposit credits are reduced and reserves accumulate. The loaning power of banks thus becomes more and more favorable as depression develops. Yet during this phase of the cycle business men do not care to avail themselves of these credit facilities. The mere fact that the banks are in a position to extend loans is not sufficient in itself to cause an

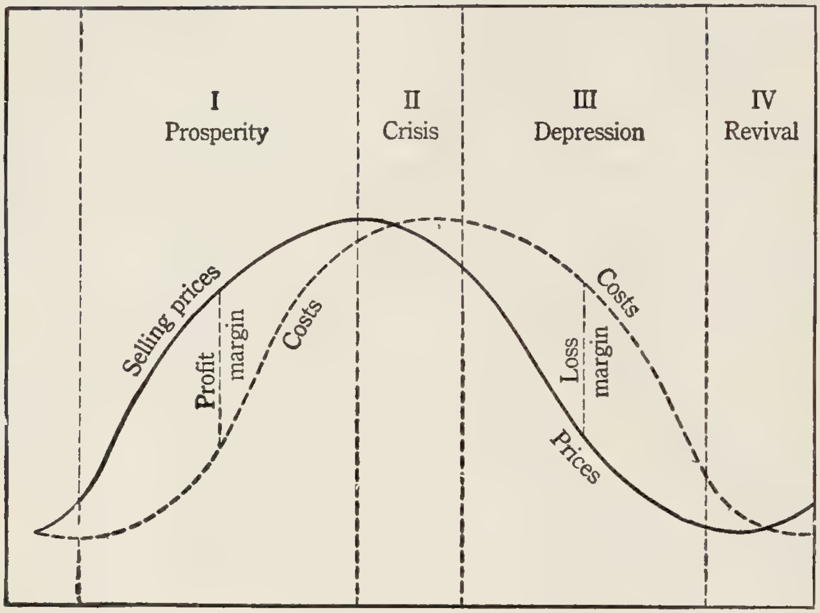

FIG. 29

Fig. 29 gives a diagrammatic representation of the four phases of the cycle which are described in the text. The diagram shows the relation between costs and selling prices in the various phases of the cycle. The space between the two indicates the margin of profit per unit of output. During the first phase of the period of prosperity the margin of profit is widening, but during the last phase of the period of prosperity the margin is gradually narrowing though still large. During the first phase of the prosperity period, therefore, the business outlook is steadily improving and so the prices of corporation stocks are rising. When, however, the margin of profit begins to narrow the business outlook is not so favorable and so the stock market begins to fall. Thus the stock market reaches the highest point long before prices begin to fall and long before the period of prosperity has come to a close. During the period of depression the profit margin is negative for many firms. During the last phase of the depression period, however, costs and selling prices are gradually approaching a more favorable adjustment, the future is beginning to look more hopeful, and so the stock market begins to move upward before the end of the depression period is reached.

upward movement. Business men do not care to borrow at the banks unless they see before them prospects of making profits, nor do banks wish to extend loans except under similar conditions. The maladjustments of costs and selling prices have to be righted, and the excessive stocks of retailers, wholesalers, and manufacturers have to be worked off before the market conditions again become favorable.

Finally, however, the continuous volume of purchases on the part of consumers depletes existing stocks. At last certain retailers are forced to place new orders. At first these orders are limited to immediate needs, for the retailers may still anticipate a further drop in prices. Gradually, however, as the orders accumulate from all quarters the prices begin to stiffen. Retailers, becoming convinced that the bottom of prices has been reached, place still heavier orders. Wholesalers increase their orders from manufacturers. So the movement is passed on from the retailers to the wholesalers, from the wholesalers to the manufacturers of finished products, from these in turn to the manufacturers of semifinished products in all the various stages, and from these to the producers of raw materials, lumber, iron ore, coal, etc. Transportation has been stimulated all along the line.

This increased demand for goods, reaching back from the retailers to all the preceding stages of production, finds expression in the enlarged purchasing power of all these dealers. Orders cannot be placed without having the necessary means of payment. Only a limited increase in orders can be paid for out of the savings of past income. Loans are placed at the banks to finance these purchases. During the period of depression large reserves have accumulated at the banks, and as a result banks are eager to extend loans to sound business men at low rates of interest. The loans thus made are credited to the accounts of the retailers, wholesalers, and manufacturers in the form of demand deposits. Thus the purchasing power of all these business men measured in terms of money units is increased, and with this enlarged purchasing power at their disposal fresh orders for goods are placed. Factories now begin to run at nearer to full capacity. More men are employed. Finally the competitive bidding for wage-earners results in higher wages. Thus the enlarged purchasing power at the disposal of business men is finally passed on to the consuming public. Purchases in one field stimulate business in a whole list of related industries. With a larger volume of sales and a more favorable margin of profit the future looks brighter. Business men become optimistic. They stock up with raw materials, employ more labor, and enter into contracts for the construction of additional equipment and plant. All these activities in a highly interrelated industrial society have a far-reaching effect, and business rapidly passes on into the phase of full prosperity.

We now turn to the theories that have been advanced in explanation of these fluctuations in prices and production. We shall limit our attention to some of the more important of these theories.

Rainfall and Business Cycles. William S. Jevons, an English economist of the latter half of the nineteenth century, and more recently Professor Henry L. Moore, of Columbia University, have advanced the theory that the underlying ultimate cause of business cycles is to be found in cycles of rainfall. The causal relation runs, it is said, as follows: the cycles of rainfall produce cycles of crop yields; these cycles of crop yields produce cycles in the prices of agricultural products; these products in turn furnish three fourths of the raw materials of manufacture and thus profoundly affect the profit margin of manufacturers. Moreover, a heavy crop results in heavy traffic on the railroads. This not only results in large railroad revenues but forces the railroads to order new cars and equipment, to buy additional coal and other supplies, thus stimulating business activity in related industries.

Professor Moore contends that he finds evidence of an eight-year cycle in rainfall, crop yields, prices, and production. He holds that this eight-year cycle appears in the rainfall in

Arizona, in North and South Dakota, in the Ohio valley, and in Illinois. The effect of these rainfall cycles is apparent on the growth of trees in Arizona; on the yield of crops in the Dakotas, in the United States, in the United Kingdom, and in France; on the wholesale prices in England; and finally on the production of coal and pig iron, basic industrial commodities. Moore finds that the maximum points for all these cycles are substantially similar and occur in the following years: 1882, 1890, 1898, 1906, and 1914.

That crop fluctuations affect business conditions may be taken for granted, but that they are the dominating factor in the business cycle is more doubtful. In the first place, it is not clear that the length of the business cycle is really eight years. On the contrary, the length of the cycle has varied considerably during the last century. In the second place, the production of crops appears to oscillate back and forth without any well-defined cyclical movements such as would correspond closely to the business cycle. In short, the evidence as to the causal relation between rainfall and the business cycle is not sufficient to be convincing.

While economists generally do not have much confidence in the rainfall theory, it is certainly clear that oscillations of crop yield have important effects on the business cycle. A good harvest may initiate or intensify a period of prosperity. If timed properly it may hasten the coming of prosperity or postpone the period of depression. Similarly, a bad harvest will intensify depression.

Overproduction and Underconsumption. One of the oldest theories of the causes of depression, and one which still persists, is that of overproduction and underconsumption. The modern representative of this theory is John A. Hobson, a British economist. He argues that the ultimate cause of these cycles is to be found in maladjustments in the distribution of wealth. With the unequal distribution of wealth that now obtains there is an overaccumulation of savings relative to the volume of consumption. Wealthy people with very large incomes are unable or at least unwilling to spend a sufficient proportion of their incomes on consumers' goods to take the current production of such goods off the market. No matter how much production is directed toward the things that would appeal to the rich, whose wants, even the most ostentatious ones, are so amply provided for, they will inevitably invest too large a proportion of their incomes in capital equipment intended for further production. Since they themselves are unwilling or unable to spend the proper proportion of their incomes on consumers' goods, the investment made in new equipment must necessarily result in the production of goods intended for the masses. But the great bulk of the population with their limited incomes are unable to purchase all these goods. Thus production outruns consumption. The inevitable result is a congested market which leads to lower prices, diminished profits, and smaller incomes for the very rich. These reduced incomes force the wealthy classes to spend a greater proportion of their incomes to maintain their accustomed scale of living. Savings are thus reduced and less is invested in new capital goods. Consumption is thus given a chance to catch up with the productive capacity of the industrial equipment, surplus goods are worked off, and prices again rise. With the return of higher prices, profits are again increased, a disproportionate share of the national income again goes to the wealthy classes, and an excessive proportion of these surplus incomes is again saved and invested in new capital equipment which, of course, must be used finally in the production of consumers' goods. The masses, with disproportionately small incomes, are unable to purchase these goods, and so production again outruns consumption. Excessively large incomes are therefore responsible for the incessant attempt to employ capital in excess of the demand of the ultimate consumer.

According to Hobson, one of two things must be done. Either the rich will have to abandon their excessive saving and resort to greater spending, or else these surplus incomes must somehow be reduced and greater equality of incomes attained. These latter ends could be obtained by the taxation of large incomes and by raising wages. A smaller proportion of the national income would thus go to the rich, a smaller proportion of the national income would be saved, and thus a balance would be restored between consumption and production, between demand and supply.

It is to be noted that Mr. Hobson is not opposed to the accumulation of capital per se. He is opposed to saving a disproportionately large *percentage* of the national income; for that compels production to outrun consumption and results in these periodic stoppages, thus defeating its own ends. His point is that these excessive accumulations of capital, instead of increasing the total national income actually diminish it, because of the unemployment that ensues. If the distribution of income were more equal and the consuming power of society thereby enlarged, production would, it is claimed, become continuous. The total national income would then be greater; and even though a smaller *proportion* of the income were saved, the absolute volume of savings and capital accumulation might well be greater than now. Hobson, of course, is not opposed to capital accumulation. He wishes to have the accumulation of capital so nicely adjusted to the total national income that the consuming power of the community will always be able to take off the goods produced.

There are many known facts which appear to support this theory. It is well known that a large part of the savings of the community come out of the corporate surpluses retained in the business during the period of prosperity. Again, it is well recognized that almost every industry is overbuilt — the steel industry, the coal mines, and many other industries. There are many industries that could not possibly dispose of their products at a profit if they continued to produce at full capacity. This would appear to give support to the notion that our capital equipment is too extensive in proportion to our national income. However, it is not difficult to account for this overexpansion of fixed capital. For one thing, when capital equipment is built it is necessary to make adequate provision for further growth and development. This factor alone accounts for much of the excess equipment. It would be uneconomical for any corporation to build a new factory just large enough to take care of its present business without making any provision for expansion. Moreover, it is exceedingly difficult to forecast the future accurately, and many errors of judgment are made. Part of the excess capacity can be justified by the prospects of growth, but a part must be charged up to unwarranted optimism. Moreover, a shift in demand caused by the introduction of new products may leave an old industry, which faces a declining demand, oversupplied with fixed capital. Again, an industry may shift from one region to another because of a combination of circumstances. Thus the textile industry is at present developing rapidly in the South while mill towns in New England are declining, and flour mills are expanding in Buffalo while those in Minneapolis are operating at greatly reduced capacity. Furthermore, new inventions make old equipment obsolete and account for a considerable amount of unused capacity. Finally, many industries are overbuilt in order to carry the peak load made necessary by the fact that we have a business cycle. If the business cycle could be eliminated, capital would be forced out of the overdeveloped industries and the consequent increased volume of capital seeking investment would cause interest rates to fall. Enterprises could then be started that cannot now be developed because the rate of interest is too high to make them profitable. It is not difficult to think of many schemes which are technically possible but which cannot earn the rate of interest that capital now commands and which therefore remain undeveloped — water-power plants, reforestation, drainage schemes, irrigation, railroads, canals, houses, and so on interminably. There is no end to the possible uses of capital if it could only be obtained at a sufficiently low rate.

The fatal objection to the overproduction or oversavings theory is the fact of the persistence of a relatively high interest rate on long-time investments. If people persist in saving, as Hobson has it, there is nothing to prevent the interest rate from falling until it becomes a minus quantity. The mere fact that the rate of interest remains high is proof conclusive that the volume of savings is not sufficient to satisfy society's demand for savings. That these savings are at present invested in industrial plants which cannot continuously produce at capacity can be adequately accounted for on the ground indicated above.[1]

Money and Credit. R. G. Hawtrey, a British economist, contends that trade fluctuations are due to disturbances in the available stock of money and bank credit. A contraction of currency occasions the slackening of demand for all classes of commodities, the fall of prices, the lack of employment, the shrinkage of profits, and the low rate of interest which are the concomitants of a trade depression. On the other hand, an increase of the currency leads to a stimulation of demand, a rise of prices, a high demand for labor, inflated profits, and a high rate of interest, all of which are the symptoms of active trade. A depression in trade is in essence a general slackening of the money demand for commodities; an expansion of trade is a general augmentation of the money demand for commodities. An expansion of trade occurs when the amount of credit money in existence is less than the bankers think prudent, having regard to their holdings of cash, and so they lower the rate of interest to encourage borrowing. Both dealers and producers borrow more money from the banks, credit money is increased, prices rise, wages are increased, cash is drawn out of the banks with which to pay wages, and bank reserves fall. At this point the bankers conclude that the amount of credit money in existence is more than they think prudent, having regard to their holdings of cash, and so they raise the rate of interest to reduce the excess. The rise in the rate of interest increases the cost of holding stocks of commodities, and dealers accordingly proceed to reduce their stocks by giving fewer orders to the producers. Production is therefore cut down, and wage-earners are thrown out of work.

---

[1] See Problem 3 at the close of this chapter for a further criticism of Hobson's theory.

Now what causes this expansion and contraction of currency? Hawtrey's answer is that in the banking institutions which prevail in the world as it is, there is an inherent tendency toward fluctuations. When credit money is curtailed, then business is depressed, unemployment ensues, and wages fall. Now the most important item in the community's demand for cash is the wages bill. If wages fell immediately as soon as banks began to curtail the credit expansion, cash would flow into the banks, the curtailment of credit would quickly be checked, and a stable condition might then be reached; but, instead, wages stay up for a considerable period, and when they eventually do fall and cash begins to flow back to the banks, it is found that the credit contraction has been excessive since there now develops an overaccumulation of cash reserves. Eventually the banks must get rid of this excess, and the result is another period of inflation. Similarly, when the banks expand their loans, wages for a time fail to rise, and therefore cash is not drawn from the banks. Eventually when wages do rise and cash is withdrawn from the banks to meet the wages bill, the banks find that they have overexpanded, that their reserves are unduly low, and so curtailment again becomes necessary. Thus the cycle perpetuates itself.

Self-generating Cycles. Professor Wesley C. Mitchell, the director of research, National Bureau of Economic Research, holds that the cycle is self-generating. Once started (and perfect stability in a complex society with a money and exchange economy is scarcely conceivable), it tends to perpetuate itself. Mitchell says that the cycle is not due to the influence of disturbing causes from the outside, but to processes which run regularly within the world of business itself. Prosperity becomes intense and breeds a crisis, a crisis runs on into depression, depression after a while begets a revival of activity, and the revival grows into prosperity. Thus the business cycles run an unceasing round, each cycle growing out of its predecessor and merging with its successor. For the explanation of business revival and prosperity it is not necessary to look to the occurrence of some propitious event, such as a good crop, or the development of new industries (for example, the electrical or chemical industries), or a favorable piece of legislation, or an increase in the export demand. "The quiet processes of business recuperation during dull times are quite competent to develop into revival without the adventitious help of any disturbing circumstances."

Why does one phase of the cycle generate subsequent phases? The answer is to be found in part in the accumulation, or storage, of stocks of goods. In the period of prosperity and rising prices, stocks of goods are piled up for two reasons: (1) when prices are rising, retailers, wholesalers, and manufacturers all have a tendency to increase their stocks before prices rise still more; (2) in periods of prosperity the demand for goods is brisk, and so it becomes desirable to carry larger stocks to insure an adequate supply. But of course retailers, wholesalers, and manufacturers cannot continue to *increase* their stocks indefinitely. As soon as they cease *increasing* their stocks (even though they continue to maintain them at the high level reached) the demand for goods at once declines in all markets except consumers' markets. Retailers buy less heavily from wholesalers, wholesalers less heavily from manufacturers, and manufacturers less heavily from the producers of semifinished and raw materials. Production is thus slowed down, wage-earners are thrown out of work, and this in turn reduces the demand in the consumers' markets. Prosperity lasts as long as stocks are piled up, but this movement eventually produces a crisis and finally depression. At this point consumers' purchases begin to fall off and retailers become anxious to *reduce* their stocks. They therefore buy sparingly from wholesalers, supplying the consumers' demand from their accumulated stocks. Similarly, wholesalers will wish to reduce their stocks, and so they buy less from manufacturers than they sell to retailers. In like manner manufacturers will purchase less from the producers of semifinished goods and raw materials than they sell to wholesalers and retailers, since in large part they too are able to supply such orders as come in from their accumulated stocks. Thus the falling off in demand is increasingly greater the farther the market is removed from the ultimate consumer. Eventually the stocks are all used up; but in spite of depression and unemployment consumers continue to purchase at a fairly even pace, partly by utilizing personal credit at the stores and partly by drawing upon savings accumulated in the period of prosperity. Finally the stocks are largely used up. Retailers are now compelled to increase their purchases from wholesalers, and these in turn from manufacturers. In consequence of these increased purchases and the improved market conditions, dealers and manufacturers again begin to accumulate stocks. Retailers buy more from wholesalers and manufacturers than they sell to consumers, and manufacturers in turn buy more from the producers of semifinished goods and raw materials than they sell to wholesalers and retailers. Production is increased; more wage-earners are employed; and so the consumers' demand is reënforced and strengthened, and this demand in turn stimulates still more the demand of retailers, wholesalers, and manufacturers. Thus the alternate accumulation and using up of stocks keeps the cycle moving forward from one phase to the next.

Price Fluctuations and the Business Cycle. Professor Irving Fisher has advanced the view that the fluctuations in the purchasing power of money largely predetermine the fluctuations in trade. It is generally recognized that a rising-price level stimulates trade and that a falling-price level depresses trade. In a rising-price period profits are increased (as we have already explained) because expenses, such as interest, rent, salaries, and wages, lag behind the rise in prices. Moreover, rising prices engender optimism, encourage speculation, and induce retailers, wholesalers, and manufacturers to stock up not only because they anticipate a good market but also because they anticipate a further rise in prices. As a result production is stimulated and industry expands. Falling prices, on the other hand, wipe out profits, force all dealers to reduce their stocks lest they be caught by a further fall in prices, and reduce the purchases of retailers and wholesalers; thus production is curtailed. Rising prices act as a stimulant on trade; falling prices depress trade. Fisher therefore contends that "changes in the price level afford an almost complete explanation of fluctuations in the volume of trade." This rise and fall in the price level, in turn, is controlled by extraneous forces, such as government inflation policies, banking policies, legislation, foreign trade, gold-mining, crops, and many other factors.

Fisher denies, then, that the cycle is self-generating. He holds that Mitchell's view overlooks "friction." If we pull a twig and let it snap back, we set up a swaying movement back and forth. But the twig, once deflected and then left to itself, soon stops swaying. Friction brings it to rest. So in business we must assume that the effect of any initial disturbance would soon wear off after a very few oscillations of rapidly diminishing amplitude. The resultant business cycle would speedily cease altogether if dependent only on its own reactions. To keep it up, there must be applied some outside force.

Progress and the Business Cycle. The distinguished Swedish economist Gustav Cassel also holds the view that the business cycle is not self-generating, that there is a tendency toward a "gradual balancing of the strength of the upward and downward movements as a result of the restrictive reactions." Thus a state of equilibrium would be reached if things did not happen from time to time which cause a renewal of prosperity. In the first place, we have technical progress. The railroads and the electrical industry are examples of new technical advances that opened up new and remunerative uses of fixed capital and stimulated their construction. The business cycle, Cassel contends, is essentially a variation in the production of fixed capital. Another event which has the same effect as inventions is the opening up of new countries. Finally, a country with a rapidly growing population demands a corresponding increase in fixed capital. Once the cycle is started, the oscillations back and forth recur in diminishing waves until the movement finally comes to rest. For a limited time, therefore, the principle of action and reaction produces a set of recurring waves. The reactions set up are as follows: the new invention or the new resource raises the remuneration derived from fixed capital, the production of fixed capital is stimulated, loans are applied for at the banks to finance the expansion. This means a diversion of the community's purchasing power in the direction of capital goods, the prices of capital goods rise, wages lag behind, a wide margin of profits is created, a larger proportion of the national income is saved, and since this is used to buy capital goods the prices of these rise still more. Finally, however, the prices of raw materials and wages rise, the margin of profit is reduced, savings are reduced, the means of financing the capital expansion is curtailed, interest rates rise, the prices of capital goods fall, and thus the production of fixed capital declines. This shows itself particularly in building construction. Thus depression automatically follows from prosperity. But depression brings out forces in opposition to itself, — low prices of material and low wages. These factors cheapen the production of fixed capital. At the same time depression results in a low rate of interest, and this in turn raises the prices of securities. The depression period is therefore one in which it is profitable to build new plant and equipment and sell them to the public in the form of securities at a profitable figure. Thus the low interest rates, the low prices of materials, and low wages put an end to depression and bring on a new advance. But eventually these forces play themselves out and a state of equilibrium is reached, or rather would be if it were not for new discoveries, new resources, and the growth of population.

In this view the business cycle is an inevitable phenomenon in a progressive society. The forces of progress — the discovery of new methods, new products, new resources — arouse bright hopes, engender optimism, and cause more fixed capital to be built. There results an expansion of credit and of purchasing power, a rise in prices, a maladjustment of costs and selling prices, a disproportionate production of capital goods compared with consumers' goods, and finally depression as a result of these disturbed relations. Thus the business cycle is held to be a function of the forces of progress and would exist, so Cassel contends, even in a progressive socialist society.

It will be noted that these factors of progress tend to develop price fluctuations. But then, even though we should be able to stabilize the value of money and thus do away with the fluctuation of the general price level, we should still have the business cycle to some degree, it is argued, so long as we live in a progressive society. Inventions and new discoveries would still disturb the equilibrium between the production of fixed capital and consumers' goods. And even though the general price level were stabilized, there would be no assurance that maladjustment would not occur between the prices of consumers' goods and producers' goods, cost prices and selling prices, agricultural prices and industrial prices, the price of raw materials and finished products. If the general price level were stabilized, the business cycle would certainly be minimized, but (so Cassel thinks, at least) the business cycle will always exist so long as we have spurts of inventions and discoveries that in turn result in the extraordinary development and expansion of fixed capital.

The Business Cycle an Unsolved Problem. It is not possible in the present state of economic knowledge, to give a thoroughly satisfactory explanation of the causes of the business cycle. To some extent it remains an unsolved riddle. Many economists are studying the problem, and it is one of the most fascinating though perplexing fields in economic theory. The apparent disagreement among economists is, however, not so great as at first appears. The literature on the subject is highly controversial, but the difference between the various writers is to a large degree one of emphasis.

There is one point to which we wish to call attention with respect to these theories. It should not be forgotten that a description of the business cycle is a record of the actions of human beings. The business cycle is, of course, a product of human behavior. When it is said, for example, that the business cycle is a function of the money economy, what is really meant is that the money economy presents certain environmental stimuli to human behavior which cause actions such as are found in the various phases of the cycle. The same is true of the rainfall-cycle theory. All these theories are really trying to explain a certain aspect of human behavior as we find it in the modern economic order.

Stabilization. No matter what the causes of the business cycle may be, it is agreed by practically all writers that certain influences may be brought to bear upon business men that will tend to check their excesses of optimism or pessimism, as the case may be.

There is ground for believing that something can be done to minimize the effects of the cycle through dissemination of more complete knowledge regarding the current movements of prices, production, stock of goods on hand, unfilled orders, etc., so that business men may more wisely manage their affairs with a view to the general condition of industry. During the last few years there have sprung up many forecasting agencies. The Departments of Commerce and Labor of the Federal government have done notable work recently in disseminating statistical information with respect to business conditions. Trade associations, as we have previously shown, have been legalized, and these give much valuable information to their membership. All this information will help to minimize errors of judgment and can scarcely fail to mitigate in some measure trade fluctuations.

A more radical proposal is to stabilize the price level.

Toward this end several proposals have been made. The one made by Professor Irving Fisher, of Yale University, has already been discussed in the preceding chapter. There we had in mind, however, chiefly the elimination of the long-run secular trends of prices. So far as the business cycle is concerned, his method would operate through the control of reserves. Increasing or decreasing the weight of the gold dollar (since gold is now used almost exclusively as reserve money) would automatically have the effect of increasing or

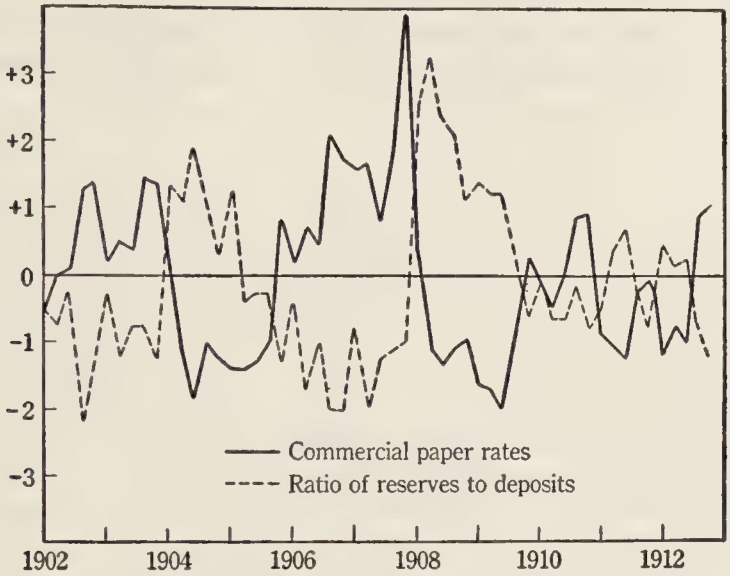

FIG. 30

Fig. 30 shows that when bank reserves are low in relation to the deposit liabilities outstanding the banks protect themselves by raising the discount rate. Thus the rate of discount fluctuates inversely with the reserve ratio, falling with a high reserve ratio and rising with a low reserve ratio. The movements of the business cycle can be plainly seen in these fluctuations. During the period of prosperity the banks are expanding their deposit liabilities, the reserve ratio is falling and thus the discount rate rises, reaching the crest of the wave in the crisis period. Then, in the period of depression, loans and deposits are scaled down, and the cash which was in active circulation during the period of prosperity is returned to the vaults of the banks. Thus the reserve ratio rises rapidly with a corresponding fall in the discount rate. This is the period of depression. With a high reserve ratio and low discount rates the money market is favorable to a revival of business, which, with the gradual adjustment of costs to selling prices and the depletion of stocks, finally comes into evidence decreasing the reserves of the banks. Under this system bankers would continue to be guided in their discount policy by the size of their reserves, as they always have in the past. But the reserves themselves would be regulated with a view to stabilizing the price level.

The proposal of Professor O. M. W. Sprague, of Harvard University, is to abandon the reserve ratio as a criterion for guiding discount policies. In the past discount rates have been automatically regulated by the reserve ratio. Professor Sprague proposes to control consciously the discount rate with a view to regulating the price level. When, for example, factories are running at full capacity and labor is fully employed, it is clear that any further expansion of loans and bank credit can only result in an inflation in prices. At this point Professor Sprague would raise the discount rate, no matter how abundant the reserves of the banks may be. In short, he would

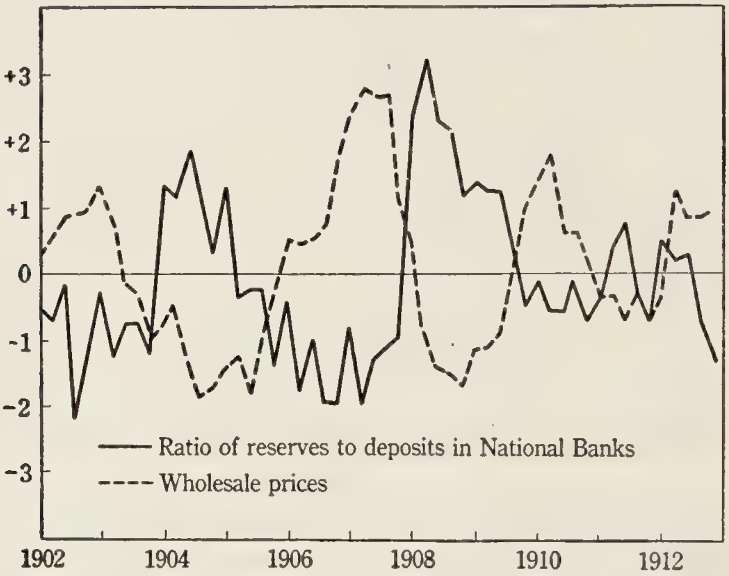

FIG. 31

Fig. 31 shows the pre-war relation of the reserve ratio to wholesale prices. The fall in the reserve ratio betokens an expansion of bank credit and rise in prices. The period of depression, on the other hand, brings a fall in prices, a contraction of bank credit and so a rise in the ratio of reserves to deposits adjust the volume of circulating media to the needs of industry and commerce in such a way that prices would be stabilized, rather than allow the ebb and flow of bank credit to run its course with the attendant upheaval in prices.

It is to be noted that Professor Sprague's device is not intended to abolish the limitations imposed upon the lending power of banks by our legal-reserve requirements. Even though the legal-reserve requirements were abolished, banks would still have to limit their loans if their legal reserves were low, in order to maintain their solvency. If commerce and industry are rapidly expanding and the money lawful for reserve is not increasing, it would therefore be impossible under his system to increase the circulating media of the country sufficiently to maintain a stable price level. His device only operates to check inflation when the legal reserves are excessive. It is at this point that the superiority of Fisher's method becomes apparent; for his device permits an adjustment of the circulating media to the needs of industry and commerce regardless of whether the natural trend of prices is up or down. It may be charged, however, that Fisher's method applies the brakes too slowly.

Since the middle of 1922 the price level has been more or less stabilized round the index number of 150 compared to the pre-war base. The discount policy of the Federal Reserve Banks during the last six years has been a potent factor in bringing about this result. In Fig. 32 we have superimposed on the curve of wholesale prices the curve of discount rates for the New York Federal Reserve Bank. The high discount rates that prevailed in 1923 brought falling prices, and the rapid reduction of rates in the early part of 1924 resulted in a sharp upswing of prices. The high discount rates from November, 1925, to July, 1927, brought declining prices, and the reduction of the discount rate in the last half of 1927 resulted in an upward movement of prices.

The effect of an increase or a decrease in the rate of discount is in part psychological. If the rate of discount is raised, business men are deterred from undertaking new building construction and dealers are reluctant to increase or even to maintain their holdings of stocks of goods, orders are curtailed, bank loans are reduced, and the volume of monetary purchasing power declines. On the other hand, when the discount rates are lowered, optimism is engendered, constructional work is carried forward, new enterprises are undertaken, dealers increase their stocks, and trade is stimulated.

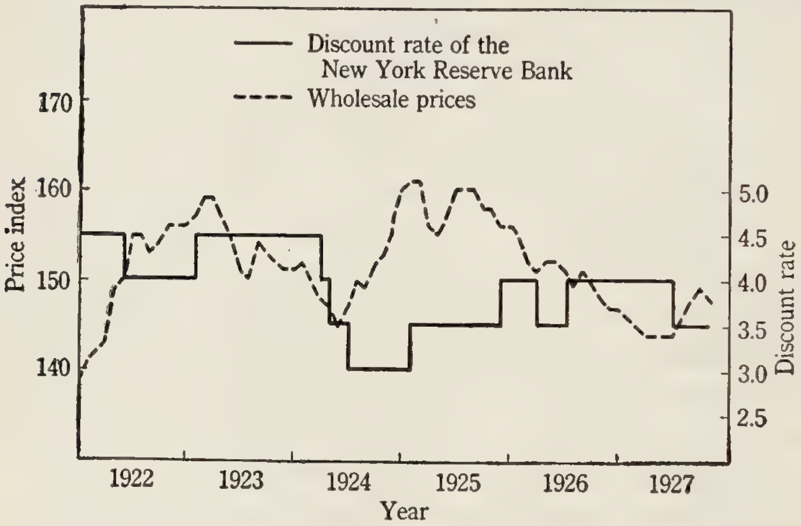

FIG. 32

Under the Federal Reserve system it is possible to apply still another brake to inflationary movements. The Federal Reserve Banks may unload their holdings of securities on the open market. The effect of such sales is to tighten up the money market, since in consequence of these transactions funds are transferred from the member banks to the Federal Reserve Banks. On the other hand, when prices are falling the purchase of securities by the Federal Reserve Banks transfers funds to the member banks and to the community at large. The member banks are thus enabled to lend more freely, purchasing is stimulated, and a check is imposed upon the declining volume of trade.

An illustration of such actions is found in 1924, when the Federal Reserve Banks increased their holdings of goverment securities about fivefold from February to November. The effect was to ease up the money market, to stimulate bank loans, and to raise the level of wholesale prices. The extent of the rise in prices may be noted by referring to Fig. 32. Similarly, the Federal Reserve Banks purchased a large volume of government securities from May, 1927, to the close of the year, and this action was followed by an upswing of prices.

Another proposal for stabilization would require all governmental units, and especially the Federal government, to plan the building of public works over a long period of time in such a manner that these undertakings could be pushed forward vigorously as soon as a decline in trade begins to appear. It is especially desirable that these public works should be financed out of additional bank loans so that the monetary purchasing power of the community may be increased.

Value of Elimination of the Business Cycle. Is it really desirable to eliminate the business cycle? The business cycle is a dynamic force in the economic order. In the period of prosperity young men and new industries have a favorable chance to establish themselves. The prosperity phase of the cycle serves as a ladder enabling new aspirants to climb into the entrepreneurial class. Moreover, the depression period serves to squeeze out the inefficient establishments and drives business men, under the pressure of dire necessity, to eliminate waste and improve their management and technique.

There are certain classes that make a living from speculation, and it is clear that the fluctuations of the business cycle increase the possibilities in this line of activity. These groups are naturally opposed to the elimination of the cycle. But many employers, perhaps most, would be glad to eliminate the uncertainties and worries that the cycle engenders. They wish to make their income from the production of goods, not from the uncertainties that the business cycle necessarily entails. Moreover, they would like to free themselves from the labor unrest which springs from price fluctuations, from wages lagging behind the cost of living, and from the inevitable wage reductions that accompany price declines. From the wage-earner's standpoint the business cycle is a serious menace, since it is the most fundamental cause of unemployment. Witness the great unemployment in 1921–1922. Especially do the unskilled suffer, since they are the first to be discharged. Employers are not likely to discharge a skilled man, because he is an important part of the organization, but laborers can readily be picked up at any time. Nor can the unskilled sustain themselves as successfully over a prolonged period of unemployment. Their wages being lower, it is harder for them to accumulate sufficient savings to tide them over. Nevertheless past experience shows that even the unskilled save a surprising proportion of the earnings received during the period of prosperity.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "There is a large and growing school of economists . . . who regard the trade cycle as essentially a monetary phenomenon, and in consequence expect to find the most efficient remedy in some modification of banking and currency policy." — BELLERBY. Do you agree? Why, or why not?

2. "The effective causes of trade fluctuations must be looked for in the sensitive and omnipresent machinery of money and credit." — HAWTREY. Argue for or against this statement.

3. Suppose a larger proportion of the national income were spent in consumers' goods and a smaller percentage saved, is there any reason to suppose that as a result there would be no fluctuations in prices? Suppose now such fluctuations occur. Would not all the phenomena of the business cycle — the lag of costs behind selling prices, the wide margin of profits, the piling up of corporate surpluses, the excessive accumulation of stocks, and finally the break in prices — at once appear? Is it likely then that the business cycle could be eliminated by more spending and less saving, by increasing the income that goes to the consuming masses, and reducing the incomes that go to the saving capitalists? Doubtless there are good grounds for desiring a more equal distribution

of the national income than now prevails, but it is very doubtful that such a program would help to reduce the severity of the business cycle.

4. "In the depression there is a low rate of interest, and this has a restorative effect upon enterprise; in the high conjuncture there is a high rate of interest, and this acts as a brake." — CASSEL

5. It should not be concluded that a country can indefinitely prolong prosperity by the simple process of continuously inflating its currency. Such a procedure may, it is true, prolong the period of prosperity for several years. For example, during the years from 1920 to the middle of 1923, while other European countries were passing through a period of great depression and unemployment, Germany, with its constantly inflating currency and rising price level, showed many of the earmarks of prosperity, despite the political stress and turmoil of the period. Factories were running at relatively full capacity and wage-earners were fully employed. However, the rapid rise in prices placed a severe obstacle in the path of trade both internal and external, for no one knew what the monetary unit was worth from one hour to the next. Moreover, the sky-rocketing of prices so profoundly disarranged the distribution of wealth and income that industry found it virtually impossible to adjust its productive processes to the changing currents of purchases. The maladjustments created by the fearful upheaval in prices became so pronounced that the final collapse, which came during the latter half of 1923, was inevitable.

BOOK IV. THE DISTRIBUTION OF WEALTH AND INCOME

PART I. FUNCTIONAL DISTRIBUTION

CHAPTER XXIV

WAGES

The Meaning of Distribution. With this chapter we begin a new division of the study of economics. In the preceding books we have examined the principles that control production, the pricing of particular goods, the value of money and the general level of prices, and the fluctuations in industry. Now we turn to the study of the distribution of wealth and income.

The term "distribution" means the sharing of wealth and income among the people of the community. It does not mean (as it often does in common usage) the physical distributing of products from the places where they are manufactured or grown to dealers or consumers. In this and the succeeding chapters we shall be concerned with the forces that determine how the wealth and income of the country are shared among the people of the country.

A moment's reflection leads us to conclude that whether wealth is equally or unequally distributed depends in part on whether income has been equally or unequally shared among the people. For although one may acquire great wealth by inheritance, the original accumulator must have enjoyed a large income. Of course wealth gives rise to income, and the process of amassing wealth is cumulative: having an income somewhat above the average, a thrifty person may accumulate wealth, and the possession of this wealth brings him a still larger income. But the fact still remains that inequalities in wealth are dependent on inequalities of income. It is logical, therefore, to treat of the distribution of income before that of wealth. The distribution of the total national income among individuals or among different groups of individuals is called the personal distribution of income. But with certain exceptions, most persons receive their incomes from the control of one or more of the four factors of production. And whether any individual receives a large or a small income depends in part on how much wealth he owns. It also depends on the amount of income that is allotted to the factors of production which he controls. If the returns to land are high, for example, landlords will receive large incomes; if wages are high, laborers will have large incomes and probably a large share in the total national income. Therefore the personal distribution of incomes depends on (1) the distribution of wealth and (2) the absolute and relative amounts of income received by each of the four factors of production. It is necessary to understand the principles that govern the shares of the four factors before we can discuss with understanding the personal distribution of incomes. That phase of the theory of distribution which deals with these principles is called functional distribution.

The Importance of Distribution. The personal distribution of wealth and income is of the greatest importance in modern society. Many writers following David Ricardo have asserted that distribution is the principal problem of economics. Although this opinion may not be held universally, there is general admission of the vital relationship between the existing inequalities of wealth and income and many of the most difficult and important controversies in economic and political affairs. A few of these relationships may be mentioned to show concretely the bearing of inequalities in distribution on some of these problems.

The control of industry is largely in the hands of those who own the land and capital used in production. The dominance of the property-owners is not, of course, complete. They are and always have been subject to governmental regulation, to influence by other forms of social control, and to limitation by the opposing forces of labor organizations, especially during the past century. But when all allowance has been made for other sources of control, it is yet correct to say that the owners of property dominate our economic life. In a democracy this situation gives rise to many political struggles. Having the power to control economic affairs, the owners of property seek to secure their position by dominating the government in matters touching property rights and property incomes, and the unpropertied classes reply by seeking to dominate it for the benefit of labor. There would still be a great deal of struggle between economic groups if wealth and incomes were equally distributed, but inequality intensifies it.

The problem of distribution has several bearings on the production of wealth. It is urged by many writers that a larger income for the laboring classes would result in greatly increased labor productivity. But the means of accomplishing the increase in labor income are difficult to discover. By some it is supposed that present inequality of distribution might easily be remedied by legislation. They assert that the possession of large amounts of wealth by the rich is the result of either present or past unequal opportunities for acquiring wealth; that if the special opportunities or privileges which are now the perquisites of a few were destroyed, the poorer classes would receive larger incomes and so could afford more efficient standards of consumption. At the opposite extreme are those who assert that inequalities of distribution are caused by inequalities in the natural abilities of the population, and they therefore deny the possibility of reducing inequality in distribution by legislation directed toward changes in the legal regulations governing economic life.

The principles of distribution are involved in many other problems and questions constantly before Congress, the state legislatures, and local authorities. The protective tariff has been held by its defenders to be the bulwark which protects American labor from the competition of the pauper labor of other countries. It is said either to cause our relatively higher wage level or to prevent its reduction. Other people assert that the recent restrictions on immigration are the first legislative acts which have actually protected American laborers from the competition of foreign laborers.

In the realm of private business a knowledge of the principles of distribution is also significant. The managers of public utilities and of other enterprises using large amounts of durable capital are constantly confronted with the question whether it is better to use labor or machinery to perform certain kinds of work. If wages are to rise in the future, some substitution of machinery for labor must take place; if they are not to rise, a heavy investment in such capital goods may prove unprofitable. A corporation which has a large bonded debt must often decide whether interest rates are likely to rise or fall in the future. If they are to rise, then it will be good policy, other things being equal, to issue long-time bonds; if they are to fall, it will be better to issue short-term obligations and refund when loans can be secured at lower rates.

The Earnings of Labor. The functional share in distribution received by human beings in return for their services is wages. The workman who is employed by another person receives a definite contractual sum which no one would fail to recognize by that name. The same is true of the earnings of all salaried employees. But it is not customary to speak of the "wages" of self-employed artisans, business men, and professional workers. Yet there is no fundamental difference between the contractual wage received by the common laborer in a factory and the fee of a physician, an architect, or a lawyer. All are payments for human service.

In many cases, however, there is great practical difficulty in distinguishing the earnings of labor from the receipts from other sources. A small storekeeper, a contractor, or a farmer always owns and employs some capital, and the last is frequently his own landlord. In all three cases the net income which the individual receives from his enterprise is a mixture of wages and capital return and, in some instances, rent and profits. In larger businesses the separation between the earnings of the entrepreneur as a laborer and his receipts from property is not so difficult. This is partly due to better accounting and partly to the fact that more than one person is likely to have an equity in the capital of the business, and as a result the managing entrepreneur will receive a stipulated salary in addition to whatever returns he may get as part owner of the business.

Wages the Price of Labor. Labor receives a payment for services just as the seller of any commodity receives a price for the things he offers to buyers. In some instances the service is purchased directly by the consumer, and the wage in that case is determined according to the principles of value explained in Chapters VII–IX. When the supply of domestic servants is very large their wages fall, the demand for them remaining the same; if the demand increases, their wages rise. The same statement applies to all the professions. The principle of demand and supply controls the earnings.

There are, however, considerable differences between wages and other prices. If a merchant cannot sell his goods at the price he has set on them, he is likely to mark them down; but if a laborer cannot get what he thinks is a proper wage or fee, ✓ he may remain idle for a long time rather than accept work at a lower price. Moreover, although it is a daily occurrence for buyers to offer a lower price for a good than the seller asks — either because they expect to bargain a little or because they do not think the good is worth what the seller asks — there ✓ is less tendency for the buyers of direct labor services to attempt to bargain with the seller. If one thinks a physician's charges are too high, he usually seeks out another whose rates are lower. When labor is employed by large firms, the employer sets a wage and the workers either accept it or look elsewhere for work; or if they do decide to bargain they will in a majority of cases do so collectively through a laborers' organization.

Another difference between commodity prices and labor prices is the relative effect of custom. The prices of many sorts of goods respond slowly to changes in demand and supply because of the fact that when a certain price has been paid for some time, neither buyers nor sellers readily think of offering or asking anything different. But in the case of wages this tendency is much more marked. No doubt the tendency is due in part to an ethical notion on the part of both employers and workers that the latter are entitled to the wage they have received in the past because it is presumably a just and proper wage.

But in spite of all differences wages are still the prices paid for labor, and, like other prices, they must be explained in terms of demand and supply. The wages of those who render direct services are determined, as has been said, by the number of workers offering their work and by the demand prices of the buyers, and the demand prices of the buyers are determined according to the principles of marginal utility. But in the case of workers who are demanded along with various kinds of capital, land, and entrepreneurship to make a common product the determination of demand is more complicated. It is a relatively easy matter for one to decide whether he wishes to employ an expert to repair his motor car at a dollar and a half an hour: he has only to consult his own satisfaction and his own means to find out whether the results will justify that expenditure; but it is far from simple for the manager of a plant to determine whether a given class of workers is worth thirty dollars a week. Their product is not easily isolated, as is the product of the expert mechanic. The remainder of this chapter will be principally concerned with the explanation of the demand for labor which does not work independently, which has no separate product, but is one of several agents employed by entrepreneurs to work together to produce a common product.

The Marginal Productivity of Labor. The typical employer's demand for labor depends upon the marginal productivity of labor. The laborer works with capital and land under the direction of the entrepreneur-employer. Now the employer produces goods to sell, and he tries to produce them as cheaply as possible. To do this he is constantly on the watch for methods of reducing his costs. One method of cost reduction is to substitute cheaper agents of production for those that are more expensive. And in general he will select that combination of the agents of production which will bring his product to market at the lowest average unit cost. If wages are relatively low while capital costs are high, he will substitute labor for capital and increase his employment of workers; if, on the other hand, wages are high, he will try to put machines in the place of human hands and employ fewer workers.

Machine
vs
Man

We may suppose that the employer begins by adding labor to a given amount of capital and land. As he adds unit after unit of labor the total returns from the factors will increase. At first, if the amount of capital and land is very great, the percentage increase in the total returns will be greater than the percentage increase in the number of labor units employed; that is, an increase of 10 per cent in the number of laborers will be accompanied by an increase of more than 10 per cent in the product. But long before the combination approaches the point of lowest unit cost the added product resulting from another laborer will have begun to decline. With wages at a given figure, the employer will add labor until the last unit just pays for itself. By the phrase "just pays for itself," we mean that the added product from the last unit of labor will sell in the market for just enough to pay the wage required to obtain that laborer's services. Here we may revert to the table on page 68 and the accompanying diagram. Labor as the variable factor will be increased to the point where the marginal cost of the product equals the selling price.

If we pass from the demand of a single employer to the demand of all the employers in society, the same explanation will hold. All employers will add labor up to the point where the last additional laborer adds only enough to compensate them for his wages; further than this they will not go. If all the workers who want jobs are not taken on when this point has been reached, there must be unemployment. Those who have no work will compete with those who have, and wages will tend to fall. The price of labor will continue to go down until all the workers can be employed.[1] But, as we have seen, this point will be reached when the marginal workers throughout industry add just enough to the common product of labor, land, capital, and entrepreneurship to warrant their employment at the going wage.

N.B. To complete the explanation we may assume that when the last available worker has been employed the wage paid him is less than the added product which accompanies his accession to the combination; that is, he and all other laborers are paid less than their marginal product. Under this condition the employers would individually fall short of the highest-profit combination; hence any one of them could increase the gains from his fixed factors of production if he raised wages a little and by this means added to his labor force. But this is a resource open to all employers, and all would use it. Competition would therefore force wages up to the point where the last unit of labor received a wage equal to labor's marginal product.

In Fig. 33 the productivity of labor is represented by the curve *AC*. On the axis *OX* are measured the units of labor which are combined with the total amount of land, capital, and entrepreneurship in the nation. If *ON* units of labor are applied to the land, capital, and entrepreneurship in industry, the product of the marginal unit of labor will be *NP*. And since the units are interchangeable, every other unit of labor has the same marginal product. Each produces *NP* and each receives as its share, or wage, the money equivalent of *NP*. If the number of labor units increases to *ON′*, the marginal product of labor will decline to *N′P′*, and wages will become the value equivalent of that product; if, on the other hand, the number of labor units falls to something less than *ON*, the marginal product will rise, and wages will be increased by a corresponding amount.

---

[1] For considerable periods of time workers may resist wage reductions even though unemployment prevails.

This statement of the general theory of wages necessarily makes certain assumptions. It deals, in the first place, with "labor units" and not with individual human beings as such. If one man can do twice as much physical work in a day as another man, he is obviously putting forth twice as many "labor units" as the other. It is labor units, of course, that give rise to product, and it is the marginal labor unit that creates the marginal product of labor. This distinction between laborers and the work, or labor units, they put forth is of importance for more than logical precision. Laborers, as we know, vary considerably in their ability to do the same sort of work and also in their ability to do different sorts of work. Herein lies one of the chief reasons for individual as well as regional differences in wages.

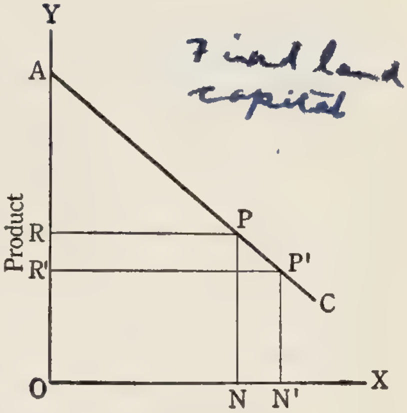

FIG. 33

Fixed land and
capital

In the foregoing explanation of the determination of wages we have also assumed free competition. We have supposed that workers compete with workers, that employers compete with employers. Trade unions have been assumed to be absent, and understandings among employers have been left out of the picture. But even if we remove these assumptions the marginal analysis will not be overthrown. Results will, however, be different from what they are under competitive distribution. If trade unions limit the supply of workers in an employment, they may be able to force wages above the marginal productivity of the normal supply of labor within the industry. But they cannot hold them at this level unless the marginal productivity of labor has also been driven up through the limitation of the supply of labor. If they exclude laborers who might desire to enter the industry and who would find employment there if no union existed, then, by their restrictive organization, they increase the supply of workers in employments outside the union. And wages will be lower in those outside employments than they would be if no union existed.

We have also assumed a perfect market. Laborers are supposed to know where they can sell their services to the best advantage; they are supposed to be free to move from place to place, from trade to trade, and from job to job. If they do not have this freedom (as in many cases they do not), then we shall find that laborers are too numerous in some places, trades, and jobs and too few in others. As a result the marginal productivity of workers of equal efficiency will be unlike and wages will be unequal. But in every case wages will approximate the productivity of the people employed.

It is sometimes said that an entrepreneur cannot discover what a laborer is worth to him, and that he knows only that he must pay the going wage; if his costs exceed the selling price he will fail in business and if they are less he will make a profit. The complete answer to this objection must be deferred in part to the next section; here we need only to point out that the entrepreneur *must* discover whether an additional laborer is worth his wage. To assert that he cannot discover when he has too much or too little labor leads to the absurd conclusion that any number of laborers might be employed with any amount of capital and land. It may be readily conceded that the adjustment of labor at the margin is inexact; only in pure theory are adjustments of any sort precise. Yet the inexactness of the entrepreneur's measurements is no reason for denying the operation of the general principle.

The device by which he discovers the productivity of labor is the common one of adding or subtracting a little labor and then noting the result. If the addition of this small amount of labor is accompanied by an increase in the product which is as great as the wage bill, then the laborer is worth hiring; if the increase is greater than the wage paid, then still more labor should be employed.

The Principle of Substitution. The entrepreneur does not, of course, carry on the experiment precisely as we have described it. In the first place, he may not regard labor in general as a factor of production. He may think in terms of different grades of labor and regard each grade as a separate agent of production. Instead of subtracting or adding labor to a fixed quantity of capital, he may add one grade of labor to a given quantity of capital and other labor. Moreover, he will probably contrast the profitableness of adding more labor with the profitableness of buying a machine, which amounts to combining more capital with the same amount of labor. If he finds that it is more profitable to add a machine instead of a number of laborers, then he will not hire additional labor; but it would not have been possible to add the machine and still improve the combination unless he had had too much labor for his capital equipment before the machine was purchased. In other words, at the rate of wages prevailing the marginal product of labor was below the wage. By substituting capital for labor the entrepreneur raises the marginal product of labor. Entrepreneurs are constantly engaged in substituting the agents of production for each other: putting more land with the same amount of labor or putting more labor with the same amount of land; putting more capital with the same amount of labor or putting more labor with the same amount of capital. And any entrepreneur tries to employ labor up to the point where its marginal productivity and the wage he must pay are in agreement.

Substitution such as we have been considering has certain limitations. The capital equipment in an established factory may not be susceptible of much modification. Machines are constructed to be run by a certain number of men, and the entire layout of the plant may be such that additional laborers can be accommodated only within limits. The same limitation applies to the combination of new machines with old equipment. But a much wider range of substitution is possible in new factories. Here the machinery and the layout of the plant may be designed to economize the use of labor, as compared with old plants, or to economize the use of capital equipment. If new plants can make better combinations of the factors of production than can be made by adjustments within old plants, then the tendency will be in the direction of the abandonment of old plants.

If only a few entrepreneurs were engaged in the experiment of finding out what is the best combination of the factors, and if experimentation and substitution were unknown throughout the rest of industry, then we should be forced to admit that in these few plants the going rate of wages determined how many laborers it would be profitable to employ with a given quantity of land and capital. And since the marginal productivity of labor depends (its efficiency being given) on the number of units employed, it would be necessary to admit that the rate of wages determined the marginal productivity of labor. And this is substantially true for any single plant. The employer cannot appreciably affect the rate of wages by discharging men or by hiring more. The number he will combine with his capital depends in part on the rate of wages; but, for industry at large, wages do not determine marginal productivity. Laborers must work to live, and, barring unemployment caused by maladjustments, sickness, and strikes, the entire labor force tends to be taken into industry. The total supply does not depend on the wages paid (at least not directly), and so the marginal productivity of labor does not depend on the wages paid.

To prove this statement let us suppose that entrepreneurs
N. generally find that unit costs can be reduced by substituting

machines for labor. Many laborers will find themselves without jobs. Wages will then tend to fall until all capable, willing workers are employed. An equilibrium will be reached when wages have been forced down to the point where labor's marginal product is equal to its wage. One may, however, ask whether it would be possible to substitute capital for labor even after this point had been reached. Might not a machine costing $100 turn out work worth $110 while labor costing $100 turned out only a $100 product? If this were the fact, the makers of the machine would quickly discover it and raise the price of the machine to $110. Or if the price could not be increased, more machines would be employed, their marginal productivity would fall, laborers would be discharged, and their wages would fall until a machine costing $100 turned out a $100 product while labor costing, say, $95, turned out a $95 product. It is evident, of course, that some labor would be required to produce the machine.

The Supply of Labor. The number of laborers is one of the 1
conditions determining labor's marginal productivity; the
others are (1) the amounts of the other factors available for 2
use with labor and (2) the ability and capacity of the laborers. 3
Here we are concerned only with the supply of labor.

The supply of labor, like that of every other factor of production except land, responds to the demand for it. If the demand for machines increases, more will be produced; if the demand for savings (new capital) grows, the amount of savings will be augmented; so, too, with labor. But the response is not the same as it is in the case of the machines: the machines will be increased until their marginal cost of production equals the gains that marginal entrepreneurs can derive from them. Applying the same reasoning to the supply of labor, we should conclude that any increase in wages would result in a further growth of population to the point where wages just covered the cost of rearing and training new laborers. This was essentially the position taken by Malthus and his early followers, who proposed the subsistence theory of wages.

Stripped of its historical setting and of its connotations in relation to problems other than wages, the argument of Malthus ran as follows: the biological possibilities of growth in population are such that in a very short time any given population is capable of doubling its numbers. In new countries, where this maximum is most nearly approximated, populations have been known to increase 100 per cent in twenty-five years. Because of the sex instinct the population *tends* to increase at this rate. That it does not always do so is partly due to certain noneconomic checks, such as war and disease; but the principal reason, according to the Malthusian theory of population, is the lack of subsistence (that is, real wages). When wages increase we may expect population also to increase until wages fall again to the point where they are just sufficient to maintain a stationary number of people.

If the argument had stopped at this point, it would have stated a cost-of-production theory of wages. It was noted, however, that among certain classes the desire of parents to give their children a good start in life and their fear of being unable to maintain their own standard of living sometimes limited the size of families. Yet in spite of this exception the reasoning led to the conclusion that the cost of subsistence governed the supply of labor and therefore wages. Since the time of Malthus the theory of population has been modified chiefly by placing greater emphasis on the voluntary control of the birth rate and on the effect of the noneconomic factors that determine it. Among the nations of the Far East it is perhaps true, even at the present time, that population keeps close to the maximum permitted by the requirements for subsistence; on the other hand, among the peoples of western Europe and among their descendants in other parts of the earth other checks than want of subsistence are at work to limit numbers.

Checks to Pop'l'n

First among these checks is the desire of parents to secure for their children a better economic position than they themselves enjoy. To accomplish this objective children must be better cared for in infancy so that they will enter the strenuous years of life with better physiques; they must be educated, and they must be equipped either with capital or with special training for a trade or a profession. But all these advantages cost money, and consequently size of the family tends to decline.

In the second place, change in social ideals has made it possible for many women to pursue well-paid trades and professions. They are thus able to secure a comfortable income and a satisfactory social position without marrying. The number remaining unmarried through life has not as yet increased to very large proportions; but the age at marriage has increased, and consequently the number of children that can result from each marriage has declined. 2

A third factor affecting the birth rate is the standard of consumption. Coincident with rising incomes has come the demand for more and more luxuries; and this rise in the level of expenditures, it is supposed, has also caused postponement of marriage and limitation of families. 3

As yet these conditions have proved most effective in limiting the birth rate among the middle and upper classes, but they are becoming more and more effective among the lower-income groups. Nevertheless there is still much truth in the Malthusian doctrine that the supply of labor varies with the demand for it. It increases when improvements and discoveries make living conditions easier, and one of the results of the industrial progress of the nineteenth century has been a great increase in the population of all the countries of western Europe and of the two Americas.

Fig. 34 shows the growth of population in England and Wales, France, and the United States, from 1800 to the present time. The use of the logarithmic scale enables us to see at a glance the *rate* at which population accumulated in the three countries. The greatest rapidity of growth has been shown by the United States, and the least by France. Yet at the beginning of the nineteenth century even France showed

Increase pop. in

the same tendencies toward increasing population as are found in England and the United States. Probably the most important cause of this increase — an increase common to all western Europe as well as to the countries for which the data are given — was the increased demand for labor. Two conditions were at work to increase the productivity of labor. In the first place, the inventions which characterized the Industrial Revolution made each worker's labor much more effective. This led to an increase of goods, easier living conditions for the entire population (though not necessarily higher real wages for the factory workers), and a higher rate of increase in numbers. As might have been inferred from the Malthusian theory, population increased because the means of living increased — not only food, but other necessities of life as well.

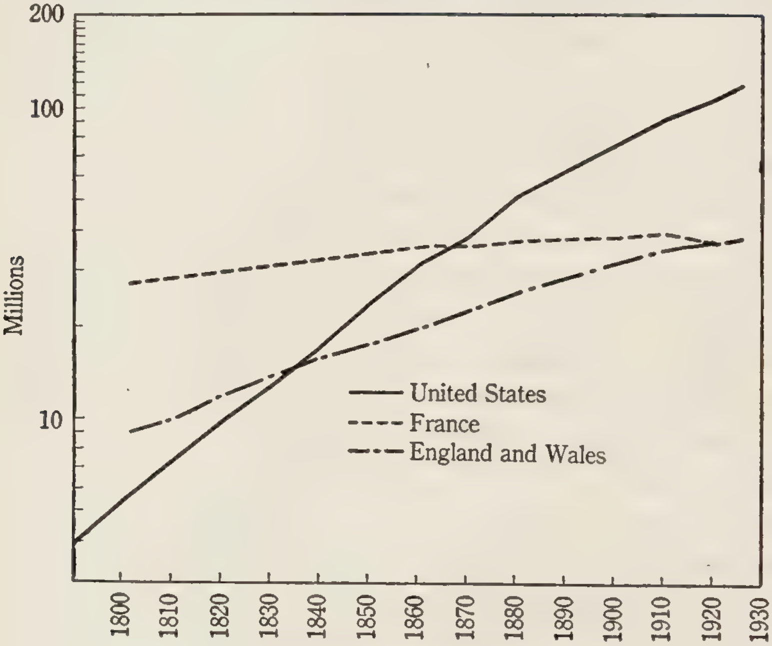

FIG. 34. Population growth

The second factor that stimulated rapid additions to the population was the opening of the fertile farm lands of the Americas and the discovery on those continents of rich mineral deposits. The coöperation of the railroad and the steamship made these natural resources of the Western Hemisphere available for the use of the people of Europe. When transportation became cheap, first England and to a less extent France, then Germany and other countries, began to import large quantities of food and raw materials, which the early development of transportation in America delivered cheaply in the Old World; and by cheap transportation the finished commodities made from the raw materials were dispatched to the most distant parts of the globe.

England led in the movement. She became a nation largely specialized in manufacturing, for the pursuit of which she had advantages in the possession of iron and coal, a large merchant marine to bring in raw materials and to send out finished commodities, a long experience in manufacturing, and a large colonial empire to serve both as a market for manufactured goods and as a source of materials. In France specialization was not carried so far. The population tended to remain more agricultural, and in manufacturing small-scale plants devoted to the production of specialties grew up instead of large mills, ironworks, and shipyards. With her larger area France might have had a much larger population than England, even without England's natural resources, if her people had turned early to mass production; but since the people retained their attachment to agriculture and to the production of the finer sorts of manufactured products, and since their failure to secure an early start put them at a disadvantage in comparison with England, France reached the limit of growth in population before England.

In the United States we see the most rapid increase of population among the great nations. Both the natural increase of numbers and immigration have contributed to this growth. But the reason for both sources of additional numbers is the same: high labor productivity due to bountiful resources and to the adoption of machine methods and large-scale production.

In the century and a quarter of growth in population shown by the chart we see at least a partial vindication of the Malthusian theory. When the supply of consumable goods became more plentiful population did increase, as Malthus had predicted. And it increased most in the country where the additional supplies of commodities were most readily available, and in those which were best fitted by institutions and by past experience to reap the gains of the new techniques of production that the Industrial Revolution made possible. In France, where the new techniques were not so easily adopted, the increase of population was less.

The rigid application of the Malthusian theory of population led to what has been called the subsistence theory of wages. The working classes, it was said by the advocates of this theory, naturally increase to the point where their wages are just sufficient to enable them to live and rear enough children to keep the population constant. If real wages increased, then the numbers of the workers would also increase and wages would decline to the level at which the family income would be only sufficient to maintain the population at a level of consumption which was little above the necessities of life. This pessimistic theory condemned the majority of the population to a hopeless round of toil and dull surroundings. Its principal error is found in the overemphasis on the tendency of the working classes to destroy for themselves all gains from improvements in technique by increasing the population. Like many early theories of economics this one failed to take account of the complexity of the forces governing the supply of labor and the efficiency of the individual worker. There is no doubt that population can increase rapidly enough to offset any gains in marginal productivity that may come from accumulations of capital, the discovery of new resources, and improvements; but whether it will do so depends on all the forces that govern the growth of numbers. And, as we have seen, the supply of food and other necessities of life is only one of the external conditions that affect additions to the population.

Another error in the subsistence theory was the overemphasis on the practical consequences of the law of diminishing returns. During the first half of the nineteenth century, and even later, economists who adhered to the subsistence theory often neglected the limiting conditions which must be present before diminishing returns result from the application of more and more variables to a fixed factor. With improvements in technique the fixed factor did not remain fixed, and each laborer came to have more resources with which to work, and not less. From Ricardo onward it was recognized that in manufacturing, increasing rather than decreasing returns were likely to prevail. But it was insisted that land was limited and that in the production of food diminishing returns were unescapable. Hence the increasing population would be unable to increase the supply of food except by a more intensive utilization of the limited supply of land. Real wages, — that is, the goods and services purchasable with money wages, — would inevitably fall. This reasoning overlooked two important conditions. Land was doubtless limited, but as time went on improvements in transportation made available for the people of Europe the produce of the agricultural areas of other continents. Moreover, the invention of labor-saving machinery for use on farms, the development of more prolific varieties of grains, and the general improvement of methods of agriculture both in Europe and in America greatly increased the available food supply.

Another factor neglected by the subsistence theory was the effect of rising incomes on the standard of living. When the incomes of the laborers increased the standard of living also increased and the growth of population was retarded to such an extent that the biological maximum was not even approached. The rising standard of living was accompanied by a falling birth rate, although other social conditions also contributed to this result.

The general level of wages is determined by demand for labor and the supply of labor. The demand for labor, which means the amount of labor the entrepreneurs stand ready to employ at a given wage, depends on the marginal productivity of labor. The marginal productivity of labor is determined by the supply of labor and by the supplies of the other factors with which it is combined, the state of productive technique being given. The demand for labor therefore depends in last analysis on the supplies of the other factors. The state of productive technique, or, as it is frequently called, "the state of the arts," depends on all the forces that make for the cultural development of a people. These lie, for the most part, beyond the field of economics. The supplies of laborers depend upon the forces that govern the numbers of the population. Some of these forces are strictly economic and others are not, or at least they are not entirely so. The supplies of the other factors of production — land, capital, and entrepreneurship — are determined in part by natural and unalterable facts. This is true of land in the broad meaning of the term, and it is especially true of mineral deposits. But here, again, we must not overlook the effect of improved technique, which cannot increase land area but which often increases the productivity of a given area of land. The supply of capital will be spoken of in a subsequent chapter; it is regulated in part by the necessary cost of waiting and in part by the facilities for directing savings to productive uses.

The increase in population shown in Fig. 34 was accompanied by greater productivity per worker arising from specialization, inventions, and new resources. In all the countries of western Europe there was a gain in real wages during the nineteenth century. It was greatest in those that were able to make the greatest gains in productivity relative to the growth in population. England and Wales, as we have seen, gained more in population than did France, yet wages did not fall below the wages of French workmen; in fact, the wage level of the former country probably rose more than that of the latter. This was because the relative increase in productivity was greater in Britain than it was in France. On the other hand, it has been due to the restriction of population that France has maintained her relatively high level of wages. Had her population increased at the same rate during the nineteenth century as did the population of England and Wales, it would have been impossible to maintain the wage level unless she had adopted specialization in the manufacture of goods produced by large-scale establishments. Even had she done so, she would not have been able to maintain the same rate of increase in numbers without reducing the standard of living, because she lacked resources equal to England's.

The marginal-productivity theory of wages is not a final explanation of wages: it is an analytical device for explaining how the forces of demand and supply are related in determining wages, *given the supplies of all the factors and the state of productive technique*. It also throws into clear relief the forces that do determine the course of wages through a period of time. Moreover, it must be remembered that it is only a general explanation. It deals with mass data and explains general, or typical, relationships and results. It supplies the analytical tools for the examination of any wage problem.

Improvements in Technique and Wages. Improvements in productive technique all have the same effect. They enable man to obtain from nature a greater return for his labor. In terms of Fig. 33 an improvement in the technique will raise the curve *AC* throughout its entire length. Hence, with any given supply of labor before the improvement takes place, an invention will raise the marginal product of labor, and wages will rise.

That inventions are beneficial to society has seldom been questioned; yet one sometimes hears the assertion that the substitution of machinery for laborers reduces wages in the industry where the substitution takes place. Whether or not wages will be reduced depends in part upon what we mean by "in the industry." If skilled mechanics have been employed to make a good, and if a machine is invented which enables unskilled workers (perhaps women and girls) to turn out the product, then the wages of the workers directly concerned with turning out the product will probably fall. And the skilled mechanics whose trade is taken away may be forced to accept work of a different sort, at which they will earn lower wages. This latter result need not follow. The introduction of the machine may be so gradual that the skilled hand-workers have time to learn new trades, or the rate at which they are displaced may be so gradual that the number of withdrawals necessary to maintain wages at the old level amounts only to the number who would naturally disappear because of death and old age.

The fear once expressed by laborers that machines might permanently displace labor, or at least reduce the total demand for labor and thus lower wages throughout all industry, has been shown during the past century to be without foundation. The use of machinery is merely a method of employing labor in production. A worker may attack a piece of work in one of two ways. If he employs direct methods he will use very simple tools; if he employs indirect methods he will make machines to aid him. Usually, though not always, the indirect method enables the worker to produce more goods than the direct method. If it does not it will never be employed.

It can be shown historically that the introduction of machinery has not decreased the total demand for labor. Against the possible decrease in demand at the point where a machine is substituted for hand labor must be set three new sources of increased demand. We may take as an example an improvement that displaces a simple sort of labor, say a machine for spraying house paint. When the machine is introduced, some painters may be thrown out of work. But the demand for mechanics to make a machine of this sort will be increased; this is the first new source of demand. Second, if many painters are displaced, the cost of spreading paint is also reduced, and in consequence there is a greater volume of demand for painting materials and for operators of the spraying

machines. Thus if the demand for the product in the manufacture of which the improvement occurred was very elastic,
then the fall in price resulting from the improvement would
lead to increased buying, larger output, and reëmployment of
the displaced laborers. Third, a great drop in the price of 3
the service, which may follow from the invention, will free a
portion of the consumer's income to be expended in the purchase of other goods.

The elasticity of the demand for the product of the industry in which the improvement is made has a great deal to do with its effect on employment and wages. If the demand for the product was very inelastic, and if the improvement considerably reduced the number of laborers required in the industry, this reduction would be permanent. Because of the inelasticity of the demand very little more of the commodity could be sold, even though the price was greatly reduced, and very few of the laborers ousted by the new machine would again find employment in the industry. On the other hand, if the price of the product fell a good deal, and if very little more of it was purchased, consumers would spend a much smaller amount of income in buying it. They would then increase their demands for other goods, and in the industries producing these goods the demand for labor would also be increased. But it must not be forgotten that when the workmen are displaced by the machines, they are unemployed and their incomes are cut off. Until they can be reabsorbed into other industries there will be little or no net increase in the demand for goods, and little or no gain to society at large. The displaced workers will lose what the consumers gain. Because there has been no increase in the total demand for goods the workers displaced by the labor-saving machinery will not readily find new jobs. Yet there is ample historical evidence to show that such workers are somehow reabsorbed. In England during the nineteenth century the percentage of unemployment remained practically constant in spite of the rapid introduction of machinery. There was, of course, some unemployment, but it was the result of maladjustments incident to a dynamic world. Since the percentage of unemployment did not increase we are forced to conclude that the introduction of machinery did not permanently displace labor.

It is conceivable that when an improvement is made in an industry for whose products the demand is very inelastic, the producers will not reduce prices but will retain for themselves all the gains resulting from the improvement. But this does not alter the foregoing reasoning: the displaced workmen will still lose their incomes and the producers will make corresponding gains, but for a time there will be no increase in the total demand for goods.

We must, however, find an explanation for the historical fact that the displaced workers are somehow reabsorbed. It is possible that the unemployed workers may by competitive bidding for jobs lower wages in other industries and in that manner create an inducement for employers to combine more labor with the other factors of production. But in modern societies with trade unions, unemployment insurance, and other institutions that tend to bolster up the resistance of the workers to wage reductions, this type of adjustment cannot always be made. At least it is difficult to bring it about within a short period of time and to a sufficient extent to absorb all the workers displaced by the machines. Under these conditions the reabsorption of the workers must await the development of new industries, new products, and new types of service. But these new sources of employment do not come into existence spontaneously. It requires inventive skill and new capital to develop new products, to stimulate new wants, and to get new industries going. New fields must be explored, consumers must be motivated to desire new products and those who have funds to invest must be convinced that the new projects are likely to become successful. Advertising plays a part in the adjustment and so does bank credit, which enables the entrepreneurs to put the productive forces in motion in anticipation of the product out of which the factors of production must eventually be paid. After the new industries have become productive the additional output of goods which they turn out increases the income and the real purchasing power of the entire country. Hence it comes about that the labor-saving improvements do finally increase the incomes of the laborers as well as those of other industrial classes. But it must not be forgotten that the workers who were displaced by the machinery could not have contributed to the income of the community until they were placed in a position where they could earn an income through productive labor.

Principles of Wage-Adjustments. Frequently wage rates are set by collective bargaining or by boards of arbitration. The factors usually considered in these adjustments are: (1) changes in productivity, (2) changes in the cost of living, (3) the minimum wage, (4) wage differentials, and (5) wage standardization. The last three of these factors are considered in the next chapter. Here we shall deal with the question whether wage adjustments according to the first two factors are or are not in accordance with the theory of marginal productivity.

In arbitration cases it is often argued that an increase in the productivity of an industry is a cause for higher wages only if it is due to an increase in the personal efficiency of the workers; but that if the cause is improved machinery or better management, wages should not be raised. According to the marginal productivity theory, labor should participate in the gains arising from increased productivity caused by additional quantities of the other factors or improvements in them. Of course, labor is supposed to benefit only to the extent that it becomes marginally more important. And historical evidence given in Chapter XXX shows that a given percentage rise in the per capita physical productivity of a nation is accompanied by a nearly proportional rise in wages. It is not, however, the productivity of a single group of workers that determines their wages, or the wage level, but the physical productivity of industry as a whole.

The "cost of living" principle of wage adjustment means that if no other factors are affecting wages, money wages should be varied as the cost of living varies. The marginal productivity theory deals with real wages and is in entire accord with this principle, since a failure to adjust wages to a rise in the cost of living would mean that real wages were reduced. However, the adjustment of money wages to the cost of living merely serves as a starting point from which further adjustments may be made if justified, for example, by an increase in the marginal productivity of labor.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. It is sometimes said that the laborers could increase their wages by first increasing their "standard of living"; that is, by demanding more and better goods for consumption. Is this statement correct?

2. If we define the standard of living as the sum of goods and services that people demand before they will assume the obligations of rearing a family, then, in the long run, the standard of living will affect wages. Show that this statement is in agreement with the marginal-productivity theory of wages.

3. How should you explain the fact that as the population of a new and sparsely settled country becomes denser, manufactures become cheaper? that the average productivity of the population increases up to a certain point?

4. "There is just a certain amount of work to be done — just a limited number of jobs in the world. If one man does more than his share of this work, there will be less for others to do. If he does as little as he must, then there will be more work for the others, and wages will be higher." Show the fallacy of this reasoning.

5. It has been argued at various times that if the rich should save less and consume more, wages would rise. Show by the use of the marginal-productivity theory that this reasoning is fallacious.

6. Does the substitution of capital for labor raise or lower the marginal product of the labor that remains employed? Does it raise or lower the marginal product of the capital?

7. Does the theory of marginal productivity explain the wages of those who render direct personal service, such as singers, physicians, hairdressers, and waiters? Are the wages of such workers affected at all by the quantity of land and capital available in the community?

CHAPTER XXV

WAGE DIFFERENCES

The preceding chapter dealt with the general relations between the supply of labor and the other factors of production. It was concerned with the most general conditions affecting the normal rate of wages. It did not take account, except incidentally, of the fact that the earnings of labor vary a great deal from occupation to occupation and, to a less extent, in the same occupation from one geographical area to another. The present chapter is concerned with some of the reasons for these variations and with the relations between the supply of labor and the other factors within different occupational groups.

Labor-Income Groups. It is a matter of common observation that there exist at the present time several fairly distinct groups of workers. The lowest income group includes all those who perform common labor. This group is characterized by lack of special skill and usually by a low level of general education. The male workers of this group do most of the heavy manual labor about factories and stores and in the construction of buildings, railroads, and highways. They supply the bulk of casual labor for all outdoor occupations. The women workers of this group are employed at simple household tasks, as scrubwomen, and as saleswomen in the cheaper shops. Both the men and the women of this group are employed to tend automatic machines in factories and to operate simple mechanical devices.

Not very much above the common laborers are the semiskilled. Among the male workers of this class a somewhat higher degree of specialized skill is required, or at least a familiarity with somewhat more complicated machinery; in clerical work and in salesmanship they occupy higher positions or work in better establishments. For most of these tasks either greater native ability to learn or a longer period of training is required than for common labor. Evidently the term "semiskilled" does not apply strictly to clerical work and selling; but the degree of attention, of independent self-direction, and of training among semiskilled workers in mechanical trades and in the other two lines of work is roughly comparable.

Above the semiskilled workers are the skilled mechanical trades and the upper groups of clerical workers and salesmen. A great deal more training or much higher native ability is required for these occupations than for common labor. The worker must have more ability to think for himself and he must accept more responsibility for his work. Among the male workers in this group we find the skilled craftsmen in construction and in the machine trades: expert carpenters, masons, electricians, lathe-operators, diemakers, and the like. In clerical pursuits are bookkeepers, secretaries, statistical clerks, and expert stenographers. The skilled workers also include the better-paid classes of salesmen and the engineers, firemen, and conductors on railroads. The skilled group is far less homogeneous than either the common-labor group or the semiskilled group, and the variation in wages is far greater.

The highest-income group is made up of professional people and business executives. Here the period of training is much longer than for other groups, and the average native ability probably much greater. Here are included engineers, chemists, accountants, lawyers, physicians, military officers, heads of government departments, and the managers of business concerns. To a greater extent than even among the skilled workers this group is lacking in homogeneity. Some of the practitioners of the professions are no better paid than expert artisans, and many of them receive less than the most successful salesmen. But it is among the managerial class that lack of homogeneity is most striking. Many managers of minor departments in large enterprises and many self-employed managers, such as small storekeepers, farmers, the owner-operators of garages, and small contractors, are not well paid in comparison with skilled labor. But in a majority of these cases either the amount of managerial capacity required is very small or the worker is an unsuccessful self-employed manager. The small storekeeper, for example, must spend most of his time and energy in doing about the same work as an ordinary salesman and clerk, and the typical farmer is chiefly engaged in ordinary manual labor.

Although this group overlaps a great deal in its lower ranges with the next lower one, it rises far above it. Each of the four groups contains many workers who could do as well for the present in occupations of the next lower group. Some of these are young people who are in reality apprentices learning the tasks of the higher group, and it makes little difference in which income group they are classified. A junior accountant may receive a salary that falls below the wage of a machinist, but he will receive higher wages if he makes a success in the higher group.

In considering the causes of the differences in income of these groups, three very important facts which partly offset the observed differences must be taken into account. The age at which the worker can command the average wage is higher in each succeeding group. The common laborer attains his maximum wage almost as soon as he has reached physical maturity; the semiskilled worker somewhat later, though not very much later. The skilled worker requires several years of training. Among the professions from three to six years of college education is common. A business executive often serves an apprenticeship through many years, climbing from minor positions to those requiring more responsibility and business judgment. On the other hand, the common laborer is likely to suffer a decline in his earning capacity at an earlier age than the skilled or professional worker.

But the fact still remains that the laborer can command his maximum income at an earlier time; and people generally prefer a present to a future income of the same size, and a small income which can be obtained at once to a somewhat larger one for which they must wait.

The second fact that tends to reduce the apparent differences in wages is the money expense and the real cost of gaining a place among the better-paid groups. The boy who is apprenticed to learn a skilled trade must usually have some additional means of support. The practitioner of a profession must incur heavy money expenses for his formal education. In addition to the money expense of training, most of the better-paid occupations require greater concentration and more attention to the work in hand than do the tasks of common labor. Hence the real cost of learning them is greater.

In the third place, the average person believes that he is much more certain to succeed in an occupation which is followed by many people and in which the pay is relatively low than in a higher occupational group. The risk of failure is much greater for the average man or woman, regardless of formal training, the higher the income paid by the work to be done. And this risk deters those who prefer a smaller but certain income from attempting to reach a higher group, in spite of the fact that they may have the native capacity to succeed in it.

In the fourth place, many occupations have advantages other than the money wage the worker receives. The agreeableness or disagreeableness of the task is a factor in the determination of the supply of labor for any given occupation. The agreeable types tend to become overcrowded, and the wages of workers in them fall below the level of the wages of the disagreeable. It must be noted, however, that when disagreeableness is coupled with low requirements as to native ability and training, wages may still be low. None of these conditions which we are considering operates independently, and none is a sole determinant of differences in wages.

But after all allowances have been made for conditions that tend to offset differences in money wages, there still remains a great deal of actual variation in the labor incomes of different groups of laborers. It is to these differences that we shall now turn our attention.

The Proximate Cause of Differences in Wages. The differences in labor incomes which we have been considering are almost entirely due to relative differences in the supplies of labor for the different kinds of work. The price paid for human services, like any other price, depends on demand and supply. The wages of common labor are low because, relative to the demand for such work, the supply is large, and because the supply is large the marginal productivity of this kind of labor is low. Capital and land are but passive instruments of production, and the former in its primary form of purchasing power has a high degree of adaptability. Land also can be adapted to different uses. Neither will be used in any industry unless it yields as great a return as could be obtained from it in alternative employments. So we find that the capital and the land used by common labor must each yield the same return as that used by highly paid labor. But since common labor is relatively plentiful, the prices of the products to which it contributes its productive energy will be low in comparison with those which are the products of skilled labor; or if the products are the resultants of both common and skilled labor, the price that must be paid for the former will be low in comparison with the price paid for the latter.

Where the different types of workers are all required to bring a product to the consumers' market common labor becomes one of many agents in production. Skilled labor, professional work, and managerial work, as well as capital and land, are the other agents with which common labor is combined. Since this grade of labor is plentiful in comparison with the other agents, its marginal product is low, and its earnings are low in comparison with the earnings of the other agents.

The salaries of professional men and managers are sometimes said to be high because these workers accept heavy responsibilities. It is said that the president of a large corporation receives a large salary because his work is very important — because his decisions involve large amounts of money. But responsibilities appear important to us when there are few who can successfully assume them. A pilot who guides a modern steamship into port is responsible for a very valuable piece of property and for the lives of many people, yet his wage is not high in comparison with the salaries of those who hold managerial positions in the steamship company; and the principal reason is that there are many people who readily learn the pilot's trade, but few who can or do qualify for the managerial positions.

Differences in wages are also said to be due to differences in the efficiency of workers. This is a correct explanation of variations within the same trade or income group, but it does not explain the variations between groups. If all common laborers were by an act of magic transformed overnight into supermen who could do four times as much work of the same sort as they now perform, their wages would be lower per unit of energy put forth. Each unit of work would have lower marginal productivity, assuming that the supplies of the other agents of production remained the same as they now are. This follows from the law of diminishing returns. Individual earnings per worker per year might be increased after industry had been adjusted to the change; but whether these earnings remained higher would depend upon the maintenance of a relatively small supply of common labor in comparison with the other agents of production.

Fundamental Causes of Differences in Wages. An explanation of the differences in labor incomes among different groups of workers must discover the forces that regulate the relative supplies of workers of the different classes. The emphasis must be placed on the relative numbers in any group and not on the absolute numbers. In the first place, laborers may be plentiful or scarce relative to the demand for the product, and in consequence their marginal productivity may be either low or high; in the second place, the laborers of a given occupational level, say common labor, may be numerous or few relative to the other laborers of different groups whose coöperation is necessary for the production of a given commodity. A good may have a high price, and yet the common labor engaged in making it may be low because it is plentiful relative to the skilled labor required for making the good. As has frequently been observed, different grades of labor do not constitute one common fund from which may be drawn at any time any number one pleases of laborers of different sorts. Each group is a partly independent agent of production whose supply is governed by somewhat different conditions. Hence common labor is combined by the entrepreneur with skilled and managerial labor in the same way that labor in general is combined with capital and land. And if common labor in any industry is plentiful while the other groups are scarce, the marginal productivity of the former will be low and its wages will be low in comparison with the latter.

One of the fundamental causes of disparity in the labor earnings of individuals and of groups is the differences between individuals in native ability. Even under uniform external conditions there are only a small number of people who are able to excel in any kind of competition, a much larger number who can attain to moderate proficiency, and a small number who can never get beyond the poorest sort of performance. Hence we should expect that where competition is present only a few people will be found at the lower levels of any sort of work, and that only a few can qualify for the more difficult tasks. But, as has frequently been observed, the distribution of human ability is not closely comparable with the distribution of labor incomes.

There are several conditions that account for the disproportionate numbers of the low-capacity, low-wage groups. An individual can attain to the higher groups if he is equipped with more than ordinary native ability, if he has been carefully trained, and if his opportunities for learning the better-paying sorts of work are good. These three — native ability, training, and opportunity — may be combined in almost any proportions. The individual of very high native ability frequently overcomes apparently insuperable obstacles, and one of no more than average ability may by thorough training and excellent opportunities go further than another who has greater inborn ability.

Training and opportunity give the children of the skilled artisan better chances of becoming skilled artisans, of entering the professions, or of obtaining managerial positions in business than the children of common laborers. The children of the professional and managerial group have still better chances. And while it may be generally correct to say that true genius is difficult to stifle, it is not correct to say that the height in the economic scale to which an individual will rise during his lifetime is not affected by the conditions under which he was reared, the extent of his education, and the opportunities open to him to learn something more than a routine, manual sort of work.

There is a tendency for the different income groups to be self-perpetuating. The children of laborers tend to remain laborers, because they inherit a relatively low capacity for anything but simple tasks, because their early training and formal education have not been such as to fit them for anything but that sort of work, and because they have fewer opportunities to come in contact with those in business and the professions who might take an interest in them and provide them with opportunities for the better-paying positions. On the other hand, the children of the upper income groups have doubtless somewhat better inherited capacity on the average, they are better cared for in infancy and so have stronger and healthier bodies, the period of their formal education is generally longer and better adapted to fitting them to hold good positions, and their opportunities for finding openings that lead to these positions are better.

The population of any country thus tends to be divided into noncompeting groups or into economic classes. The barriers between these classes are not impassable, however, and certain institutions and economic conditions greatly diminish their effectiveness. First in importance among these institutions is the publicly supported school system and the public provision of other educational facilities. Public supervision of health and public control of sanitation are also factors. Still another factor is the increase in the general level of incomes. If, as the annual income of a nation increases absolutely, the incomes of common and semiskilled laborers also increase, some, at least, of the disadvantages of the children of the laborer's family will disappear.

Differences in Wages between Different Places. Everyone is aware of the fact that the same sort of work is not paid at the same rate in different parts of the country. Very often, however, these differences are more apparent than real, even when money wages alone are considered. A trade may go by the same name, and to the casual observer the workmen who follow it in different places may seem to be doing exactly the same sort of work; but on closer investigation one often finds that there are variations in the skill required or in the conditions of work that easily account for variations in the rate of pay. And if we look about us we shall find the same variations within a relatively small area. On the other hand, there are many actual variations. In Table XVI are shown the hourly earnings of male workers in certain of the leading occupations in the manufacture of cotton goods in 1924.

The cotton-manufacturing industry is one that extends along practically the entire Atlantic seaboard. It is a highly standardized industry as far as labor and tasks are concerned. It is also a large-scale industry. Here, if anywhere, we should expect to encounter uniformity of wages. It is evident, however, that there are differences in money earnings per hour ranging from 70 to over 100 per cent as between the high-wage and the low-wage states. There is an almost regular decline in

TABLE XVI. HOURLY EARNINGS IN FIVE LEADING OCCUPATIONS IN THE MANUFACTURE OF COTTON GOODS, 1924 [1]

STATES	OCCUPATIONS				
	Doffers	Spinners	Speeder-Tenders	Loom-Fixers	Weavers
New Hampshire	$0.459	$0.480	$0.515	$0.705	$0.532
Massachusetts	.472	.507	.565	.696	.543
Rhode Island	.470	.511	.495	.695	.542
Connecticut	.369	.519	.515	.673	.522
New York	.466	.503	.562	.680	.551
Virginia	.362	.355	.470	.558	.503
North Carolina	.310	.213	.371	.457	.401
South Carolina	.262	.200	.302	.391	.328
Georgia	.283	.227	.306	.363	.314
Alabama	.255	.188	.284	.396	.298

wages as we go from New England to the South. The differences in money wages indicated by the table are offset in part by differences in the cost of living. The workers in New England pay higher prices for housing, and some other necessities than those of the South, but real wages per hour are higher in the North.

How are these differences to be explained? They can be accounted for in part by the absence of competing industries in the South. There is less capital seeking workers, and the productivity of laborers in industry is lower. The greater strictness of labor laws in the North and the firmer hold of trade unions in that district are also factors.

These differences in real wages would be greatly reduced if there were complete mobility of labor between the two sections. But obstacles to freedom of movement are numerous. The money cost of moving is a factor which keeps laborers in a low-wage district when they know that the rate of pay is higher elsewhere. Often they do not know definitely how great the differences really are. Ignorance of the labor market prevents them from selling their services in the highest market. They do not like to give up associations already formed, and they fear that somehow they will not find conditions any better in another locality. All these grounds of explanation apply to other lines of manufacture as well as to the cotton trade.

---

[1] *Bulletin No. 371,* United States Bureau of Labor Statistics.

Sometimes differences in wages between localities are due to differences in the productive abilities of the workers. High wages are really "high" only when the wage payment *per unit of output is high.* To some extent such differences in productivity account for variations in all industries.

Since the wares of manufacture tend to be sold in the same markets, it is natural to suppose that the employers of poorly paid labor make unusual profits, that the labor is exploited. Often this is the case, but sometimes it is not. The concern paying low wages may be unfavorably located. Or the management may be inefficient; that is, antiquated machinery or methods may keep the marginal product of labor below the level of other districts, and the employer does not receive an unusually large profit. Variations in marginal productivity arising from variations in the efficiency of the management or from the use of obsolete equipment cannot persist unless there is also present some limitation on the freedom of movement of workers. Where limitations are not present, the dwindling of the labor supply for inefficient firms will force the wages paid by them to the same level as those paid by concerns of typical efficiency. The owners of the less efficient plants will lose money; they will not be able to cover all supplementary costs. And this is as it should be. Society cannot afford to reward equally the results of good and poor management.

When differences in wages are due to an oversupply of labor in some localities, society will gain if the laborers are redistributed; that is, it will do so if the cost of shifting laborers from one district to another does not exceed the gain in production resulting from such redistribution. When differences are due to inefficiency of management, society will gain if the inefficient managers are forced to surrender control of the factors of production.

The Wages of Women. It is notorious that women are more frequently paid extremely low wages than men and that they often appear to be paid lower wages for the same work. Some part of this difference is due to the greater physical strength and endurance of men. Another cause for the lower wages of women and girls is the fact that most of them are not permanent workers. The younger group of women, who do not remain long in industry, must go into work that can be quickly and easily learned. They take up the so-called blind-alley jobs. Young men, because they expect to be permanent wage-earners, are more likely to avoid such occupations.

But after all allowances have been made for such differences, the fact still remains that women usually receive lower wages for the same work than do men. This difference is chiefly due to the limited opportunities open to women. Workers compare the wages offered in one line with those paid in other lines. They will not accept a wage which is less than that offered by another employment. For capable, well-trained men these alternative opportunities are more numerous than for women. In the long run, if women can perform a given task as well as men, only women will be employed to perform it, and the wage will be lower than if men were employed. From this difference in wages the employer derives no benefit. Competition causes the prices of products to fall, and the consumers who buy the goods receive the gain.

A final reason for the lower wages of women is found in their weaker bargaining power. Since they are largely temporary workers, and since they have less responsibility for dependents, they are not so easily organized into trade unions; they do not maintain such strong unions as men and do not fight so tenaciously for high wages.

Minimum-Wage Legislation. The problem of the exploited worker and of the worker whose efficiency is low because his wages are poor has been attacked by means of minimum-wage laws. It is not our purpose to describe the provisions of these laws. The principles according to which they operate do, however, fall within the scope of the present chapter. Two dissimilar conditions are encountered in the application of minimum-wage regulation: (1) wages in the industry to be regulated are below the wages paid for workers of equal skill and capacity in other industries; (2) wages are not below those of other industries.

If wages are lower than in other industries the employers may be taking the difference in profits, and the application of wage regulation will simply force the employers to turn over to the laborers their full marginal product.

Often wages are unfairly low, and yet the employers receive no benefits from the low labor cost. Competition among them may cause all the gains to be passed on to the consumers of the products made by the underpaid workers. Here the effect of regulation is to raise prices and drive some employers and some employees into other industries. From this transfer some social gain will result. Before the regulation took effect society was receiving too much of the products of the sort upon which the underpaid employees were engaged, and too little of some other products. The regulation tends to remedy this maladjustment of production.

Underpayment of labor may exist without any gain either to employers or to consumers. The employers may be inefficient as a group or there may be failure to eliminate the unfit through competition. The inefficient employers cut costs by reducing wages instead of increasing efficiency. Labor, being immobile and redundant in supply, cannot help itself. Regulation in this case should drive out the inefficient managers or cause them to find another means than wage reduction for cutting down their expenses.

If management is efficient, if there is no exploitation of labor, and if the industry is not overexpanded, the attempt to raise wages by law can only result in driving both labor and capital out of the industry. One effect of the raising of wages by law is to cause employers to select workers with greater care. The least efficient are dropped from pay rolls; the more capable are retained. Therefore the excluded workers will not be benefited; they may even suffer from the operation of the law. Those who are retained will gain, and the consumers of their products will lose.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. What conditions are principally responsible for the immobility of labor? Is mobility desirable?

2. Immigration into a country has the general effect of reducing wages. But a heavy immigration within a short period of time might raise the wages of *some classes of labor*. Which classes?

3. It is a common observation that wages tend to be higher in new countries than in old. How can you account for this tendency?

4. Even if we could provide equal opportunities for all workers, — equal education and equal opportunities to enter trades and professions, — there would still be differences in wages, because of differences in inherited ability. Does it follow from this statement that society should make no attempt to provide equality of opportunity? Would the national income be increased by greater equality?

5. Is it desirable that men should earn no more than women in the same occupations?

6. In old countries where social stratification is well-defined and relatively permanent, differences in wages tend to be more permanent. Why?

7. Enumerate the principal forces that tend to eliminate differences between the wages of different groups of workers.

8. If wages for all types of workers were fixed by law at a level above the competitive marginal productivity of these workers, could the law be enforced?

9. A minimum-wage law might attempt to (*a*) standardize wages for the same grade of labor within the same industry, (*b*) standardize wages for the same grade of labor in all industries, (*c*) raise the general level of wages for all grades of labor within a given industry. Which attempt is most likely to succeed? Why is it important to raise abnormally low wages to the standard wage of the trade?

CHAPTER XXVI

THE RENT OF LAND

The two primary factors of production are land and labor. In the preceding chapters we have studied the conditions that determine the wages of labor in general and of different groups of labor. We have seen how labor's share in the common product of all the factors is determined. In this chapter we shall apply the same type of analysis to the share received by the owners of land.

Land as a Factor in Production. In economics the term "land" includes all the different natural resources that man calls to his aid in production: farms, building sites, waterfalls, coal fields, oil pools, quarries, and mineral deposits. These natural resources differ among themselves in many ways, yet all have certain physical and economic characteristics in common: all are limited in amount, all are spatially immobile, all are nonreproducible. It is because of these essential similarities that we shall consider them all together and call the share which accrues to them in competitive distribution "rent."

The similarities of natural resources must not be allowed to obscure their differences. Mineral deposits are nonreproducible, just as farm land is. Unlike the latter, however, mines, quarries, and oil pools can be easily exhausted. Moreover, there is a great deal of variation among the different items of any kind of natural resource in respect to richness and accessibility. Some farming areas are extremely fertile and others are relatively barren; some mines yield nearly pure metal and others only low-grade ores. Some agricultural areas and mineral deposits are near the ocean or other bodies of navigable water; others are buried in the inaccessible interiors of the larger continents. Some are found in those sections of the tropics where an unhealthful or enervating climate makes utilization by civilized man next to impossible. Because of these conditions sometimes naturally rich portions of the surface of the earth are sparsely populated and little used, whereas areas far less fertile swarm with workers. In some instances the progress of technique has reversed the relative desirability of two areas. The first settlers in the United States could not transport goods between the Mississippi Valley and the Old World. For them the sandy Atlantic coastal plain and the rocky hills of New England were better land than the fertile fields of the interior. Later the building of canals and railroads reduced the difficulties of transportation; today we regard the Mississippi Valley as the most valuable agricultural area on the continent, and the lands of New England and most of the Atlantic coastal plain are regarded as distinctly inferior in productivity.

The Meaning of "Rent." In economics the term "rent" formerly meant a payment for the use of land as it exists or originally existed in nature. The improvements that man makes on the land were thought to have their own peculiar return. Yet some of these improvements are, relative to the life of any individual, so durable that they tend to merge with the land itself. Of such a nature are the construction of "made lands" along water fronts, the removal of stones and forests from farming areas, the drainage of swamps, and the leveling of sites for buildings. Hence in more recent writings on economics we find that rent is defined as the payment for the use of land as it originally existed, together with the payment for such man-made improvements as are, for all practical purposes, as permanent as the land itself.

In business usage there is no clear distinction between rent and the return for capital improvements on the land. It is said that a building "rents" for so many dollars a year, which is the payment the tenant makes for the use of both the land on which the building stands and the building itself. This return is a payment for (1) the rent of the land, (2) the use of the building and other improvements, such as streets and sidewalks, and (3) the time of the landlord or his agent in supervising the total investment in both land and building. In economic terminology we should say that the payment made by the tenant consisted of (1) rent, (2) interest and depreciation, (3) wages, and possibly (4) profits; and if we wished to speak of the total payment we should call it the "business rental" (or simply the "rental") of the building.

If each landowner used all the land he owned, there would be no payments of rent. Yet the land would have a value, and its use would be worth just as much as though it were all leased to other people by the owners. Rent which landlords derive from the use of their own land is called *implicit* rent; that which they receive from tenants is called *explicit* rent.

The distributive share with which we are dealing is the *normal competitive* return to land. But competition does not work with faultless precision in the market for land leases. Sometimes the advantage is with the tenant; sometimes it is with the landlord. Hence the rent called for in a leasing contract may be either more or less than the economic rent of land. Some land, especially in the cities, is leased for very long periods of time. When the future rental is agreed upon in advance, there is almost certain to be error in calculating the economic rent, and contract rent does not always agree with economic rent.

This discussion of terminology cannot omit one further distinction. Many writers speak of the "rent" of capital goods. Now it is true, as will be pointed out in the chapter on interest, that durable capital goods yield a return which is in some respects similar to land rent; but the identity is never complete, and hence it is well to limit the term "rent" to payments made for the use of land or to these and implicit land rent. When we wish to imply that the return to capital goods is like the return to land, we shall employ the term "capital rent."

The Demand for Land. The demand for land, like the demand for labor, is usually a joint demand. When it is used directly by the consumer for residence, some capital and labor are always necessary for coöperation with it. In the case of producers' use it is obvious that land is never used alone. In our study of value we found that the price which the entrepreneur will offer for a good used jointly with other goods depends on (1) the value of finished product, (2) the scarcity of the good itself, and (3) the scarcity of the coöperating factors. And this principle applies to land. A miner working a bed of minerals will offer a high price for the privilege (have a high demand price for the land) if the product is valuable, if the deposits of such minerals are scarce, and if the labor and capital used in mining are relatively plentiful and therefore cheap; if, on the contrary, the ore is low in price, if the deposits are plentiful, and if labor and capital are expensive, he will not be able to pay the landlord very much for the use of the mine.

The demand of all society for available land is affected by the same conditions. Land in a densely populated, highly industrialized community generally has a high value because many workers are seeking to use it, because the population requires a large amount of land products for subsistence and raw materials, and because there is much capital as well as many laborers available for combination with the land. And, other things being equal, the denser the population the higher the rent of the land on which the population lives.

The payment of rent by an entrepreneur depends upon the conditions of joint demand laid down in the preceding section. We may first ask the question Under what conditions would the entrepreneur pay no rent? It is clear that the mere desire of the landowner to receive rent has nothing to do with the answer. In the first place, the entrepreneur cannot be compelled to pay for the use of the land if the supply of it is so plentiful that the refusal of any particular landowner to allow his land to be used will not create a scarcity of land. So long as land of the best grade exists in such quantities that some of it is out of use all the time, no rent can arise; for the tenant upon whom the demand for rent might be made would simply substitute some of the unused tract for the one he had been occupying. And since the land is not scarce, the combination which the entrepreneur would make between land and the other factors would be such that the latter would yield the maximum average return per unit. If, however, land became so scarce that the variable factors labor, capital, and entrepreneurship were added in such quantities that diminishing returns set in, then the loss of any amount of land from the combination would force the use of the same amount of labor and capital on a smaller area of land, and the total return from the entrepreneur's operations would be decreased. The entrepreneur would then find that he could afford to pay rent for some additional land which would increase the productivity of the entire combination.

We may suppose that an entrepreneur makes a combination of 100 acres of land with 50 units of capital and 50 units of labor, and that this combination yields 1000 bushels of grain. This is the best combination. If 90 acres were used, the product would be something less, say 950 bushels. If, now, land becomes scarce, — that is, if there are some entrepreneurs without land or some who need more land to make better combinations with their labor and capital, — then bidding for the land will begin to take place. The user of the 100-acre tract will either have to give up some land or pay rent for it. If he gives up, say, 10 acres, he will find that his labor and capital will produce less grain. If the observed decrease in product is 50 bushels, then the average return for labor and capital will have declined from 10 bushels per unit to 9.5 bushels. Evidently he could as well afford to give the landlord 50 bushels for the use of the land as to reduce his farm to 90 acres, and if there is competition the landlord will be able to obtain rent for the entire tract.

In this illustration nothing has been said as to why land becomes scarce. As far as an entrepreneur is concerned, scarcity arises from the fact that other entrepreneurs are willing to bid for the land. They are willing to do so because they are in possession of labor and capital but without land, or because they have already applied more labor and capital to their holding of land than has the first entrepreneur and can therefore improve the total productivity of their farms by acquiring some additional land.

The same analysis will apply if we assume that a farmer comes from a distant place to begin operations in a well-settled territory. He will find that all the land is owned and in use. We may suppose that he will go to several farmers and ask them to allow him to cultivate a portion of the land they occupy; but if land is scarce in the community, if all of it has had applied to it that amount of labor and capital which brings maximum average returns, he will find that none of them will allow him to use any portion of their land unless he pays rent for it. The reason is, of course, that if they parted with any portion of their farms they would be forced to apply their labor and capital more intensively to the areas remaining to them, and as a result the average return per unit of these factors would fall.

The scarcity of land arises from natural limitations. And in any given stage of the development of productive technique the land resources available for a given population located in a given geographical area are for practical purposes fixed. Some land can be "made." For agricultural uses the irrigation of arid regions, the draining of swamps, and the diking of rivers and the sea may add small amounts. For urban uses the filling of areas otherwise unsuited for the support of buildings may also be resorted to. But these methods can add but a negligible fraction to our natural resources. The more significant method of increasing the supply of agricultural and building land available for the use of the population is through changes in the methods and cost of transportation. Densely populated areas, such as England, Holland, and the northern Atlantic seaboard in the United States, produce in the immediate locality neither their food supplies nor the raw materials for their manufactures: they draw on the lands of other and distant regions for these supplies. But their ability to do this depends upon a low cost of carriage. These facts are important when one considers the relation between the amounts of capital and labor on the one hand and land on the other. For many sorts of production the combination between these factors of production must be considered on a world-wide scale.

The Amount of Rent. The amount of rent that will be paid for any piece of land depends upon the marginal productivity of that land. Let us assume that a farmer is tilling a tract of 100 acres with the application of 100 units of labor and capital. His crop amounts to 1000 bushels, which sell for $2000. If he subtracts a small part of the farm, say 5 acres, and leases or sells it to another, he finds that his total product decreases. This follows from the principle of diminishing returns, since the 100 units of labor and capital are now applied to a smaller area of land. If the diminution amounts to 25 bushels, then this is the productivity of that small part of the farm. It should be noted that the labor and capital will now produce a crop worth only $1950. Obviously the farmer would not allow another to occupy this 5-acre tract unless he paid $50 annually for the privilege. The loss accompanying the subtraction of 5 acres measures roughly the productivity and the rent of that tract. The rent of the entire farm is, therefore, approximately $10 an acre.

There is, of course, a slight error in our calculations, because 5 acres is a rather large fraction of the entire 100 acres. The subtraction of so large an amount of land might very easily make the use of 100 units of labor and capital unprofitable and cause the farmer to sell or lease some of it to another person. If, however, we should suppose that the subtraction was a very small portion of the farm, then this subtraction would not affect the productivity of the labor and capital (except so far as the land was responsible for that productivity), and the difficulty would be obviated. If 1 acre was subtracted and the total crop reduced by 5 bushels, then the calculation of the rent would be substantially correct.

Farmers probably do not often experiment in exactly the way we have illustrated, but they apply the same principle. They find that rents are fixed by the market and that the prices of labor and capital are also determined in the labor and capital markets. They then make the best combination of the factors, these prices being assumed, and they discover what is the best combination by adding now a little more capital and labor and now a little more land. If typical farmers find that they can profitably dispense with some land and make a better combination by leasing smaller farms, then the demand for land will decline and rents will fall; if, on the other hand, they find that with the prices of the factors as they are, better combinations can be made by adding more land to the capital, then the demand for land will rise and rents will also rise. If all the farmers own the lands they occupy, the same effects will follow; but since no contract rent is paid, the increase in the demand for land will manifest itself in higher prices for it. Each farmer will want more land, and the prospects of bettering the combination of the factors will induce each to bid for additional acres. But the reason for this rise in the price of land is the rise in economic rent.

Rent and Differences in Land. In the earliest explanations of rent great stress was laid on the fact that different tracts of land vary greatly in their natural productiveness. The existence of differences in the productive capacity of different areas of land complicates the explanation of rent, but it does not introduce any new principle. The explanation is complicated because the supply of land is made up of competing areas which are not homogeneous. In Fig. 35 are represented three grades of land, *A*, *B*, and *C*, which are unequally productive. The variable factors labor, capital, and entrepreneurship are applied to the land in combined units and in doses of 100 units priced at $1 a unit. Since the prices which the entrepreneurs must pay for these factors remain constant, each dose can be measured either in physical units or in money. In the diagrams the physical measure of the variables is employed.

If the demand for the product of all three grades of land, say wheat, is so small that less than 130 bushels (that is, 100 × 1.3, the product of 100 variables on grade *A* when the best combination is made) is required by consumers, some of grade *A* will lie idle. Less than 100 variables will be combined with less than the total area of grade *A*. Since some of that grade is idle none of it can have any value and all land will be a free good. Wheat will sell for its marginal cost in terms of the variables alone, or 77 cents a bushel (that is, $100 ÷ 130); but when more than 130 bushels are required, more of the variables will be applied to grade *A*, marginal cost per bushel will exceed 77 cents owing to diminishing returns, grade *A* will become a scarce agent of production and all of it will yield rent.

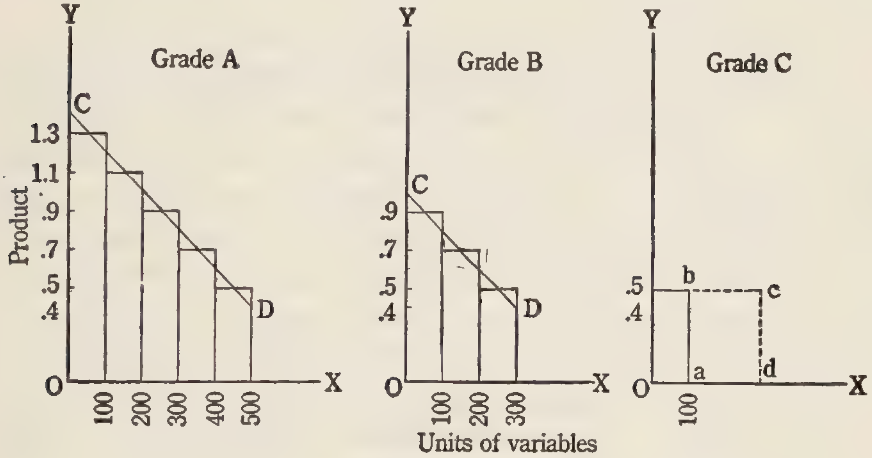

FIG. 35

But when demand has increased to the point where it is possible to apply a third dose of the variables to *A*, it is found that 100 variables will produce as much on *B* as on *A*. The total yield derivable from the third application on *A*, 90 bushels, can also be obtained by cultivating *B*, and the increase of production by intensive cultivation of the best grade will occur simultaneously with its increase by extension to land of the second grade. The two margins, the intensive on *A* and the extensive on *B*, must be equally productive if there is free competition and if the variables are mobile, since farmers would not apply more than 300 of the variables on the best land, where the productivity of the fourth application of 100 units of the variables falls to 70 bushels, until they had exhausted the possibilities of less costly production on land of the second grade.

When consumers demand 710 bushels at $2 a bushel, the best land will receive five applications of the variables, or a total of 500 units, costing $500. On land *B* 300 units will be applied. Land *C* will be the marginal land, because on it $100 worth of variables just pays for itself by yielding a crop of 50 bushels, which will sell for $100. The price of wheat will now remain at $2 a bushel until all the land of grade *C* has been cultivated to the point of maximum average returns.

We now turn to the measurement of rent on the *A* land. We find that this land now produces, with the aid of 500 units of the variables, a crop of 450 bushels, which sells for $900. How much of the crop will the landlord receive as his share? The answer is that he will be able to compel the tenants to pay over to him the full contribution of his land to the total crop of 710 bushels. This amount can be ascertained in a number of ways. Suppose that the variables were deprived of the use of land *A*. Then, in order that the demand might be met, they would have to be employed on the marginal land, where their productivity would be only 250 bushels. Hence the owner of land *A* can compel the tenants to pay to him the difference, or 200 bushels. This is the "bushel rent" of his land. The measurement of the rent of land *A* can also be found by subtracting the total cost of the variables, which measures their productivity in alternative employments, from the total value of the common product of these variables and land *A*. The total value of this product is seen to be $900, and the total cost of the variables is $500. Hence the money rent of land *A* is $400, which agrees with the amount of the bushel rent as previously computed. By similar calculations the rent of *B* land can be shown to amount to 60 bushels, or $120.

The question will probably occur to the student whether the existence of land rent is dependent upon the differences in the productive capacity of land. If we mean by this question, Are land differences as we find them a factor in determining the rent of particular farms and of particular tracts of city land? the answer must be an affirmative one. The existence of poorer land does affect the rent of the better grades, because it offers alternative opportunities for the production of goods. And the rent of any given area does depend upon its superiority to no-rent land. On the other hand, it is not true that rent would not exist if there were no differences in fertility, location, climate, and the like. Let us assume that land *A* in the illustration is the only land available for the supply of wheat. When more than 130 bushels are demanded, rent will arise on this land because the withdrawal of any part of it would reduce the total productivity of the four factors. And when the intensive margin has been extended to the point where 100 units of the variables produce only 50 bushels (demand now requiring a total of 450 bushels), the rent of the entire area of land *A* will be 200 bushels, or $400.

In the preceding discussion of the rents of *A* and *B* lands, we have assumed that the variables were applied to all grades of land in relatively large separate doses. The areas of the polygons in the figure thus measure the *average* productivity of the successive doses of the variables. In this procedure there is some error. Evidently the first units of the first dose of 100 units of variables on land *A* should be more productive than the hundredth unit. The productivity would vary from 1.4 to 1.2 bushels. And the marginal productivity of the last unit in the total of 500 units would be something less than 0.5 of a bushel. Hence the line *CD* is a more accurate representation of the theoretical operation of the principle of diminishing returns. It is also evident that if the marginal productivity of the variables applied on land *A* is assumed to be less than 0.5 of a bushel, the marginal cost of production will have to be higher than $2 a bushel, and the marginal-demand price will have to exceed $2. The rent of land *A* will also be greater than was found when doses of 100 units of variables were employed. But the principles of explanation are the same whether the doses are assumed to be very small or relatively large. In dealing with the measurement of rent on individual tracts of land, the supposition of doses of 100 units is probably nearer reality, and it does not lead to an appreciable inaccuracy; but in dealing with the rent of all the land in an entire country, 100 units would be a very minute amount compared with the total capital and labor applied to the land; hence in our explanation of the generalized theory of rent we shall use a continuous curve, such as *AC*, in Fig. 36, to represent the productivity of the variables.

The Generalized, or Social, Theory of Rent. Thus far the explanation of rent has been given in terms of individual parcels of land and in terms of the calculations of individual entrepreneurs. To explain the rent that would arise throughout the entire country, let us suppose that the land area is made up of a given number of entirely homogeneous units and that all the entrepreneurs are marginal. Let us regard land as the fixed factor, and labor, capital, and entrepreneurship as the variables. If, now, we apply the existing supply of the other factors — labor, capital, and entrepreneurship — to the fixed area of land, we shall find that as unit after unit of the former is added, the return per unit (the marginal return) diminishes. Finally, when all the variables have been applied, a certain product is obtained. In Fig. 36, *a*, this is represented, by *OAPN*. But the last unit of the variables produces only *NP* product. The sum of the variables is, then, responsible for the product *ODPN*, and the land is responsible for *DAP*. The latter area in the figure represents rent. But this illustration seems to imply that the rent is a residual share, and this is not the case. In Fig. 36, *b*, we have represented the labor, capital, and management as the fixed factor (several factors combined) and land as the variable. When *OR* units of land have been applied to the other factors, the marginal product is now *RQ* and the rent is *OSQR*, and the residual area *SAQ* is the combined product of the factors other than land. The rent of land is thus seen to be determined, as are wages, by the marginal product of the factor to which the return is paid — labor in the case of wages, and rent in the case of land.

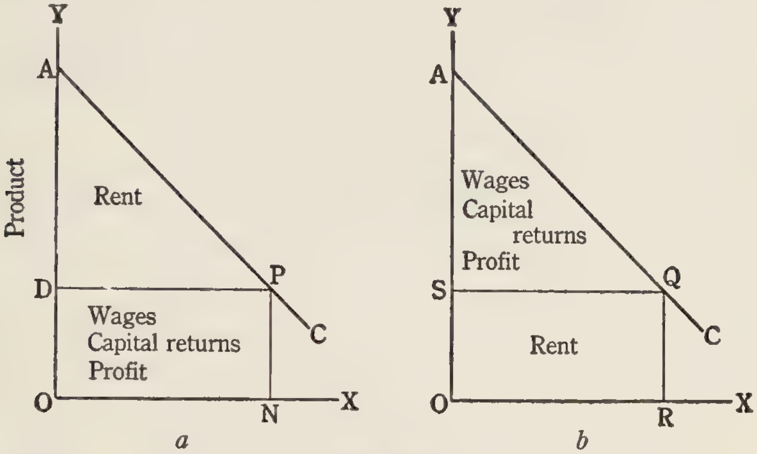

FIG. 36

There is a degree of unreality in the two preceding illustrations, but it does not arise out of the method of analysis. The lack of reality arises from the assumption that land is homogeneous and that the labor and capital of society are applied to land in "doses" that contain an unvarying combination of these factors.

The illustrations serve one important purpose: they are near enough to the truth to show what tends to happen to rents when the other factors increase in amount and the best land in a country has all been brought into use and cultivated past the point of diminishing returns. In Fig. 36, *a*, an addition of the variables past the point *N* on *OX* will result in diminished returns to these variables and in an increase in rent. It will have this effect if there is no improvement in technique, so that the curve, AC, representing diminishing returns, remains fixed. And such is the tendency. As population increases, the margin of cultivation is forced farther and farther toward the point where no additional raw materials and food can be wrung from the soil and each acre of land becomes more and more important. Its marginal productivity is increased, and the landlord's share in distribution rises.

To combat the tendency to diminishing returns, man has many weapons. But they are all of one type; they all lead to a more complete utilization of resources. We call them "inventions" or "scientific discoveries" or "more efficient labor" or "better methods of agriculture." The effect of these improvements is to raise the curve AC in Fig. 36, *a*, farther above OX. And as long as men can devise new ways of capturing natural forces or of avoiding the waste of those they now hold, population may increase without reducing the marginal product of labor and capital.

Agricultural Rent. The economic rent of any particular farm is determined according to the productivity of the land which is included in it. But this statement obscures certain details that are of great practical importance. When one says that a certain amount of labor and capital is used in connection with a tract of land, he is ordinarily understood to mean the labor and capital used directly in farming. But production is usually not complete when crops are harvested. Wheat is not a consumable good : it must be transported to central markets, milled into flour, distributed to consuming centers, and baked into bread before it is ready for consumption.

The labor and capital used in transporting, milling, and baking are just as much concerned in the production of bread as is the work done in the wheat fields. That this is the case is known to everyone. But the implications of these facts are not always so well recognized. If much labor and capital must be used to carry the wheat from remote land where it is grown to the consuming centers, the productivity of this remote land is thereby reduced. Such land adds less to the total finished product than does that which lies nearer these centers, although it may yield just as large a crop in return for a given amount of agricultural labor and capital. And since its productivity is less, its rent is less. It is customary to speak of differences in rent due to location as *situation* rent. Thus a farm near a large city which is no more fertile than one distant by a thousand miles may rent for 50 per cent more. And this difference is situation rent. All that comes of this distinction between ordinary rent or *fertility* rent and the excess return on the more favorably located land is to call attention to the fact that fertility alone does not determine how good or how poor the land in question is. The reason why it does not is that fertility alone does not determine what the marginal product of the land shall be; something depends on location.

Agricultural land yields various products, some of which are very scarce relative to the demand for them, and others are plentiful. This relative scarcity arises from the fact that some crops may be grown on almost any sort of land, whereas others can be produced with reasonable facility only on limited areas. And these limited areas have higher rents than other lands, although they may be capable of producing no more of the plentiful products than are the other lands. Lands in the citrus-growing areas of the United States yield rents greatly above the rents of the best wheat land, not because they would produce more wheat but because they will grow oranges and lemons and the wheat lands will not. Still another example is found in the relatively infertile areas near New York City, which yield high returns to labor and capital because they are especially suitable, in respect to both location and soil, for growing perishable foodstuffs, such as vegetables and small fruits.

Urban Rent. Land in a city may be used either for residence or for business. The conditions affecting the productivity of land used for these two purposes are not entirely the same. The rent of residential sites is in some respects closely akin to agricultural rent. The demand for a residential site depends on topography, drainage, and other fixed conditions affecting its natural capacity to meet human desires. But man-made conditions affect the demand for urban residential land more than they do the demand for farm land. Proximity to good schools, efficient passenger transport, public improvements, and the fashionableness of the neighborhood are all major factors in the determination of the price that people will pay for the occupancy of city land; they are less important in determining the demand for farms.

The rent of urban land used for business purposes is determined almost entirely by location. Yet the desirability of different locations is largely man-made. Nature does one very important thing: it limits space. Topography plays some part, but not so great a part as in the case of agricultural land.

High business rent arises in the centers of the retail and financial districts. Where traffic is great and where many people congregate it is possible to sell goods more rapidly than in the outskirts of the city. In other words, the limited area in the heart of the retail district has a high productivity because it is possible to transact a very great volume of trade on it. Hence the withdrawal of literally any square rod would cause a considerable reduction in total productivity. The withdrawal of the same area of land in the outskirts of the city would not affect total productivity by anything like the same amount. Because it is possible to drive so large a trade on a small centrally located area, lofty buildings are erected and much capital and labor are used in conjunction with the land. It is sometimes said that high rentals are the result of heavy improvements, but this is not the case. One often sees high buildings, erected by a foolish speculator, which do not cause the site on which they stand to have a high rent. Rather is it true that the great demand for the land causes it to be highly improved.

In any city of considerable size rather well-defined zones of use will be found. The highest rents are found in the retail and financial districts, where the concentration of people is greatest and the demand highest. Next in order come the districts of wholesale trade and light manufacturing, then the zone of heavy manufacturing and such businesses as coal yards and grain elevators. Between the business area and the better residential districts there usually exists a borderland which at any given time is devoted to residence. Present rents are likely to be low, but land value may be high because of the prospective use of the property for manufacturing or some other business. Beyond this zone lie the better residential areas, which are again differentiated according to the desirability of the sites.

Since urban rent is nearly all due to location, it follows that the land values and rents in a city are greatly affected by its transportation system. The retail district must develop where large numbers of people can conveniently reach it; the wholesale and manufacturing districts must be served by the railroads; even the residential districts depend for their desirability upon the ease with which people who live in them can reach other parts of the city.

The Unearned Increment. Land is the only factor of production whose supply is not affected by real cost. All labor is more or less irksome to the average person; at least, the amount and kind of work performed by a majority of us becomes irksome. Some part of the savings of society is made at a sacrifice of present consumption. Hence wages and interest are thought to be payments necessary to secure the present supply of labor and capital. The real cost of entrepreneurship stands in a much more doubtful position. It is doubted by some economists that any real cost is entailed by entrepreneurship. On the whole, however, the position of land is unique; there is no vestige of real cost connected with its supply. Moreover, whether real costs are involved or not, it is said that the landlord is in no wise responsible for the supply and cannot affect it. There would be just as much land for use, it is argued, if property in land were abolished tomorrow.

The rent paid to the landlord, it is concluded, is not a payment necessary to bring forth the supply of land. To this conclusion we must in the main agree. Yet there are several objections. The land exists as so much physical stuff. It is not a self-directing agent of production. Someone must plan for its use, improve it, and, if it is leased, attend to the business of finding tenants and collecting the rents. Since the landlord performs these functions, the "rent" which he receives is partly rent and partly a composite of wages, capital return, and profits. If there is any excess of returns above the fair payment for the performance of these duties, that payment is true economic rent and is the result of the productivity of the land and not of the landlord's activities. It is unearned and unnecessary so far as the productive process is concerned.

So much we may grant. But when it is argued, as it is by the advocates of the single tax, that society should proceed at once to deprive the landlords of their rents, we are not obliged to agree. In the first place, the argument does not take into account the fact that the supplies of some sorts of labor are not dependent on the wages ordinarily paid for them. It is absurd to suppose that the monetary returns received by a great surgeon or by a popular opera star are "necessary" to secure these services. To be sure, the high returns to these professional workers may attract more people into their fields than if only the wage of common labor were ever earned in them; but that is far from being an identical proposition with the statement that unless the opera star received $5000 a week the supply of such artistic talent would be reduced. Of more practical significance and therefore of more weight as an argument is the fact that a large number of people receive higher wages than are necessary to cause them to put forth their best efforts in production. Human activity is motivated by a great variety of objectives and conditions. It is fair, therefore, to ask the advocates of the single tax whether they propose to tax away all the unearned increment in wages. To this question at least a partial answer can readily be made. Our interest in high wages is humanitarian and cultural. We believe that for a majority of the people higher wages mean greater welfare. And so long as an augmentation of wages does not, in the long run, check production or cause a decline in the total productivity of all the people, we would not do anything to prevent that augmentation. Rather, we would do everything possible to bring it about. To this answer the opponents of the single tax make the following rejoinder. They say that the rent of land in the United States is very widely distributed. It is not received, as is sometimes said by overardent advocates of its confiscation, by a small idle landholding aristocracy. Several million farmers receive the bulk of agricultural rents. In the smaller cities residential rents go chiefly to people of the middle classes. The rents of business property are probably much more concentrated; and in the great cities all rents are likely to go to a smaller, wealthier group. The problem of the "unearned income from land," say the opponents of the single tax, is mostly a problem of the great cities and not of the country as a whole. Outside these cities rent is part of the income which helps to raise the standard of living of much the same group that is aided by higher wages.

Another objection to the single-tax program, one which runs in the same vein, comes from those who see in the taking of interest by the very rich another instance of "unearned income." That much of the interest now paid is not necessary to secure the present supply of capital might as well be admitted. Not everyone who receives 5 per cent on his investments would cease to save if the rate were lowered to 2 per cent. Since this is the case some interest, though by no means all, is just as unnecessary as is land rent. However, there is practical difficulty in discovering the receivers of unearned interest that does not exist in the case of rent. We can determine with a fair degree of accuracy how great the true rent of any piece of land is; we cannot determine whether the interest payment on a certain bond is more than just what was necessary to induce its owner to save the money he used to buy it.

It is not the purpose of this chapter to give categorical answers to controversial questions, but to point out some of the fundamental characteristics of rent. However, some conclusions may be stated on the topic of this section. Economic forces cause rent to exist. The institution of private property in land gives all persons the legal right to own land and thus to collect rent. As far as economic reasoning goes, we can safely conclude that rent is an income whose payment to landlords does not affect the total supply of natural resources. But it does not follow that it would be good public policy to confiscate economic rent. Whether confiscation would be in line with sound public policy can only be settled after all the factors affecting social welfare have been considered. Would it be for the best interests of society to take from the present owners of land the rents they expected to receive when they bought land? It may also be asked whether the settlement of America would have taken place as rapidly as it did if the pioneers had known in advance that they would be forbidden to profit from the rise in land values. Would the present resources of the country continue to be used and conserved as efficiently as they now are if rents were confiscated? Would private initiative in the further development of mines and water power be discouraged? These and other important social problems connected with land ownership must be carefully considered before one can arrive at an intelligent judgment as to the desirability of draining rent into the public treasury.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. An isolated frontiersman produces a scanty crop on a small clearing. He uses some capital to aid him. Under these conditions would it be possible for him to discover whether the land he uses yields any rent?

2. Criticize the following statement: "When the land in the state of X was first taken up it was very fertile. But because of careless cultivation the fertility of the soil was depleted and diminishing returns set in. Now the land is relatively unproductive, products are scarce, and rents have arisen."

3. How would an improvement in the transportation of the products of land affect the situation rent (1) of lands near the great consuming centers? (2) of those distant from such centers?

4. How should you expect an improvement that lowered the cost of production on the best grade of land to affect the rent of that land? the rent of the second grade of land, if there are more than two grades?

5. "Rent is a result and not a cause of high price." Is this statement true? Explain why, or why not.

6. "Since man did not make the land, the taking of rent by the landlord is taking something that belongs to all of us. It is a bold, bare robbery, like chattel slavery." Do you agree with this statement?

CHAPTER XXVII

INTEREST

Land and labor, the primary factors of production, yield returns that are determined by their marginal productivity. The third factor, capital, also receives a return which, like rent and wages, is a share in the common product of all the factors. This return is commonly called interest; but, as will presently be explained, not all the return to capital goods falls into the category of interest. Part is replacement of the original investment of the owner of the goods. But before we can proceed with the explanation of interest, we must first understand the nature and the function of capital.

Nature and Kinds of Capital. In business usage the term "capital" is given a variety of meanings. Sometimes it is taken to mean the net investment of the proprietors in the business they own, — that is, the worth of the enterprise less all obligations to persons outside the proprietorship. In other cases business capital means the total investment of both the proprietors and the bondholders. These two concepts have in common the idea that capital is the amount of the investment in the productive enterprise. But the worth of a business depends primarily on its earning power. If the earning power of an enterprise is so low that no one would give for it the amount of money originally invested, then the earnings of the concern and not the original investment determine its value. On the other hand, if the earnings show a return much greater than could ordinarily be obtained from the original investment, it must not be hastily concluded that the concern is worth more than the amount the investors have put into it. We must know whether the surplus earnings are likely to continue in the future; whether they are to be attributed to a specially favorable location, to patents, or to superior management. If we find that the conditions in the business are such that the surplus earnings are likely to continue in the future, we conclude that the capital value of the business exceeds the amount of money the owners have invested in it, and that the business capital exceeds the original cost of the plant.

The economist's concept is different from the business man's in several respects. In the first place, the economist is not concerned with the total investment. That may consist in part of goodwill built up by advertising or by custom, which is not capital. Nor is he concerned with the capitalized value of the business as a going concern. For economics the term "capital" means any physical agent of production other than man and natural resources.

All capital is produced by man. It consists of tangible goods, of physical objects that can be produced only by the expenditure of labor and the use of natural resources. Hence all capital costs something to produce. If a machine requires a certain amount of labor, and if a given amount of natural resources is likewise necessary, then the cost of obtaining these factors is the cost of the machine.

Capital is said to be saved. By this we mean that instead of using land and labor to produce goods nor immediately usable in consumption we choose to make a machine whose products will only be available in the future. Saving is the act of setting aside purchasing power for the payment of the expenses of making capital. It involves a decision to forgo immediate consumption and take a money income in the future instead.

Producers' capital consists of all produced means of production employed by entrepreneurs in business undertakings; consumers' capital consists of produced means of production which yield direct utilities to their users. The distinction is not made on the basis of any physical difference between the two kinds of capital but on the use to which the good is put.

A house owned by the person who lives in it is consumers' capital, and the same house used for a store or a work place is producers' capital. For the most part there is, of course, a physical difference. The greater part of producers' capital consists of machines, business buildings, railroads, and other equipment. There is, however, a more important reason for the distinction. Consumers' capital is always used for direct gratification of personal wants. Producers' capital, on the other hand, is used by entrepreneurs solely to produce a money income. Its employment is controlled by purely business motives, and it is valued for the monetary returns it will bring in. The employment of consumers' capital is affected by whim, fad, impulse, and all the other imponderable considerations that rule in the domain of consumption. In the discussion of interest we shall confine our attention to producers' capital, because it constitutes by far the major part of the sum of durable economic goods, land excluded.

It is not usual for the economist to call nondurable consumers' goods capital. Bread is not capital but "product." On the other hand, coal consumed in a factory *is* capital. Both coal and bread are perishable in the sense that they lose their utility and their value with a single act of use. The distinction between consumers' goods, which are highly perishable, and producers' goods, which, like coal, are also perishable, is not logical. Both would be called capital by one who wished to be entirely consistent. But there is a difference in that the coal consumed in a factory, if properly used, will aid in the production of some valuable physical good or salable service, whereas the bread gives rise only to the final product of all economic activity — the satisfaction of human desires. We identify the bread with the satisfactions themselves. We do not do so — cannot, in fact — with the coal, unless perchance we burn it in our homes, when it can no longer be considered producers' capital.

Kinds of Producers' Capital. Producers' capital is not homogeneous. Some forms are durable and others are perishable;

part is specialized capital and part is nonspecialized, or free, capital. A machine designed to perform a mechanical operation of a limited sort is called "specialized" capital; a building that can be used for a variety of purposes is called "free" capital. In general, most raw materials, metallic money, some machines, and some types of building are examples of free capital. The entrepreneur can quickly withdraw his free capital from an unprofitable employment, but not his specialized capital. So, too, he can withdraw his investment from perishable capital, but not from durable capital. From this difference arises the fact that old specialized capital may yield a net return which is far below the going rate of interest. It continues to do so because it cannot be shifted to any other use, and because it has not entirely ceased to be productive and the entrepreneur cannot afford to abandon it.

In Chapter XXV we used the term "mobility of labor." Capital also possesses a certain degree of mobility. Free capital has industrial mobility; that is, it can be diverted from the industry toward which it was intended to move and be directed to other uses. But much capital does not have a great deal of mobility; it must remain where it is first installed, and most must stay in the industry for which it was made. Yet we constantly read that capital has been shifted from one industry to another, as though it could be moved about with the greatest freedom. What usually happens is that investments have been shifted. Capital yields its owners a return which they may invest in a great variety of employments. A part of this return, as we have already said, is repayment of the investment. The machine wears out; but if the investment of money in it has been wise the owner will receive from it during its lifetime a sufficient amount to pay back its original cost plus interest. He may use the interest for the purchase of consumers' goods; but unless he wishes to dissipate the wealth he owns he must reinvest that portion of the earnings of the machine which amounts to its original cost. This we may call the replacement fund.

When capital shifts from one industry to another, it is the replacement fund which is being diverted. As shipping declined, capital was shifted to the textile industry. No diversion of capital goods could have occurred, of course. What did probably happen was that the owners of vessels failed to replace them and diverted the replacement fund to the purchase of factory equipment. In the same way, investments in perishable capital are constantly on the move. In one year the banks may lend large sums of money to one industry and in the next to an entirely different industry. Because bank loans are commonly used to purchase raw materials and to pay laborers, they are quickly repaid. When the demand for loans in one line of production declines, the banks seek loans in another quarter; hence the capital goods which the banks finance may be greatly changed in a very short period of time.

It should now be clear that capital is not a "sum of values": it always consists of physical objects. But these objects have value; and since they are very unlike physically, about the only way in which they can be reduced to a common denominator is in terms of value. It is for this reason that the business man when asked how great his capital is will usually reply by stating its estimated money value. But the money value of capital is no more capital than bushels are wheat.

The Return to Capital Goods. Capital, as a physical agent of production, yields a return based on its marginal productivity. That is, a machine is a factor of production as well as the labor that tends it and the land on which it stands. Like that of the other factors, its return is governed by marginal productivity. The entrepreneur discovers the productivity of capital, just as he discovers the productivity of land or labor, by varying the proportions of all three factors until he obtains the combination that will produce goods at the lowest unit cost. If marginal entrepreneurs find that they can profitably substitute labor for capital, then there will probably be a surplus of the latter; its marginal productivity and price will decline; and old capital will be revalued. With this fall in price more capital will be employed — a resubstitution will occur — and finally the total supply will be taken into industry at prices that correspond to the productivity of capital.

Unless the price paid for capital covers its cost, its supply will be contracted and its marginal productivity will rise. Hence there is a relation between the cost of capital and its earnings. But this relationship is very complex. In the first place, it should be noted that the entrepreneur makes a different calculation for the machine from the calculation he makes for labor. The product of labor is work, and for that the entrepreneur pays wages. He does not buy the laborer; he does buy the machine. Having invested a sum of money in it, he must look forward to replacing it when wear and tear and the progress of technique have rendered it unproductive. He does not need to make any such provision for labor. Moreover, he buys the services of labor and, within a short time, turns these services into goods, which he sells. But when he buys the machine he must ordinarily wait for his return and must take risks that come with decisions in respect to future events.

In buying a machine the entrepreneur must first estimate its marginal productivity — its annual earnings throughout its life of usefulness. From these earnings he must then deduct an annual amount which when accumulated during the life of the machine will return to him the original investment. If the machine does not earn this amount, it is clearly an unprofitable investment. And it must earn something more than this if the entrepreneur is to receive any interest on his savings invested in it. For example, a machine which cost $1000 and which wore out in one year would need to earn at least $1000 during the year before the entrepreneur owning it could recover his original outlay. If it earned $1060, and if the going rate of interest was at 6 per cent, then its marginal net product of $60 would exactly equal the interest charge for that time. To put it another way, a machine which would earn $1060 gross would be worth $1000.

If men were willing to exchange *present* goods (goods whose money value or consumers' service can be had *now*) for an equal amount of *future* goods of the same sort, then the producer of a capital good could sell it for the sum of its marginal gross products. And such capital goods would be multiplied until their costs and the marginal products were equal. But, as a rule, no exchange between present and future goods takes place unless the future goods have first been scaled down in value or discounted at the going rate of interest.

We may summarize the discussion of the productivity of capital thus far given as follows: capital goods, like other factors of production, earn returns based on their marginal productivity. But, unlike labor and land, the supplies of capital goods are affected by their direct money costs, and unless these costs can be defrayed from the marginal product of the capital goods they will not be produced. Partly because they are durable but chiefly because the capitalistic process of production requires time, capital goods yield their products mainly in the future. Since men do not exchange present for future goods unless the latter are discounted, capital goods must ordinarily yield returns that exceed their costs. If they do not, the labor that might have been invested in them will be diverted to the production of goods and services for more immediate use.

The Superior Productivity of Capital. It is evident that capital labors under the handicap of having its returns discounted at the going rate of interest. This raises the question why men ever employ it when they might easily use labor and land for immediate production. The answer to this question is found in the superior productivity of the roundabout, or capitalistic, method of production. The same amount of labor applied to the making of machinery will often yield many times the product that it could create if it were applied directly to natural resources. For example, a savage can grow scanty crops without the aid of a plow or any other capital. He merely scratches the soil with the first conveniently formed stone or stick that comes to hand and scatters seed. A somewhat more highly developed civilization would make use of crude implements and gain a much larger harvest. Modern agriculture utilizes tools that can be had only as a result of a high degree of development of the arts of metallurgy and engineering. Instead of setting to work to produce the crop in the shortest period of time and by the most direct methods, capitalistic production engages in many preliminary processes before any consumable product emerges. As the technique of production becomes more and more complex, these steps are multiplied in number, and the time that must elapse between the inauguration of the process and its completion becomes longer. The longer period of time is made necessary by the variety and number of the preliminary steps and by the fact that the machines and other equipment are durable. They are durable not by choice, but because there is no escape from durability. Here we must guard against a possible misapprehension. Men contrive to the limit of their powers to shorten the time required by the productive process. Factory-owners no more desire to have their savings tied up in raw materials and half-finished goods than merchants want to keep goods a long time before they sell them. Nor would they prefer a machine that turned out ten thousand articles in ten years to one that shaped the same number of articles in five years. But unfortunately such capital goods as machines cannot be quickly exhausted of their productive capacity, and as a consequence a machine-using industry must put more time between the beginning and the end of the productive process than the direct, noncapitalistic industry does.

The explanation of the superior productivity of the roundabout, machine, or capitalistic method is not difficult. Yet it is sometimes said that roundaboutness of production increases product. This is not the case. Roundaboutness is an unescapable incident of the method, but it adds nothing. The machine method is more productive, in the first place, because it enables man to capture and control natural forces that would otherwise escape him. To the savage the waterfall is not an aid in production, because he cannot utilize its energy; nor could modern man, were it not for the use of capital. In the second place, the capitalistic method enables man to accomplish certain results that are impossible for unaided labor: to tunnel mountains, to create waterfalls by means of dams, to move large weights and to move them at a speed beyond the capacity of man or of work animals. It is the superior productivity of many sorts of capital as compared with direct labor that causes capital to be used in production. But the fact that the method is roundabout does not increase the productivity. Yet some productive processes seem to require merely time for their completion. When wine ages it changes in quality and therefore in value. Fruit ripens and lumber cures. Back of these common terms, "cure" and "age," lie chemical and physical processes. And these happen in time; but it is not the passage of time nor the roundaboutness of the processes that produces the change.

Not all roundabout processes are superior in productivity. Many ways of spending labor in the making of machinery cannot be employed, because they are either not superior to labor working more directly with natural resources or they are not sufficiently superior to offset the discount arising from the phenomenon of interest. The entrepreneur selects from among all the variety of methods of using labor, both direct and indirect, those that yield the maximum returns. The rate of interest prevailing in the community and the productivity of the roundabout method, as compared with the direct method, both affect his decision.

Time Preference. We have said that men generally will not exchange a present good for a future good unless the latter is more valuable. This general statement is only typically true. We must note first that the statement does not say that an individual prefers *any* present good taken at random to a given future good. It would be absurd to suppose that he preferred a present good of low utility to a future good of high utility. The statement does say that he will usually prefer a present good to a future good of the same sort. In the second place, the statement does not take account of certain cases where desires change with the passage of time. The schoolboy has no present use for the paraphernalia of a lawyer's office. He knows, if he intends to become a lawyer, that he will require an office in the future, and he prefers to have it then and not now. In winter one does not usually want ice as much as he will want it during the coming summer. These and similar cases are modifications of the general statement. It is admitted that individual valuations change with the passage of time, and the individual knows this fact and takes it into account; it is also admitted that individual valuations change with changing environment, and this too the individual knows and takes into account.

The explanation of this customary discount on future goods is not entirely clear. Several reasons for it have been advanced by different economists, and we may note them briefly with the warning that all are largely conjectural. In the first place, the fact that men know the future to be uncertain is said to be a factor in causing the discount. Since we may not live to enjoy future goods, we set a lower value on them than on present goods. There is also the chance that conditions will change and that the future goods may not yield as much satisfaction as we now anticipate. It is also said that men, being generally optimistic, believe that they will be better provided with goods in the future than they now are, and therefore, according to the principle of diminishing utility, scale down the value of future income.

Whatever may be the cause of the discount on future goods, we are sure that it exists because the evidence lies everywhere about us. Men do not lend money except as an act of courtesy between friends or out of charity unless the lender agrees to return the original loan with something in addition. A child offered a choice between a treat today and a treat a month hence usually chooses the present enjoyment. This phenomenon of discount of the future has been called time preference by Professor Irving Fisher.[1] In his words, "It is the (percentage) excess of the present desirability of present goods over the present desirability of an equal amount of future goods." In its generalized statement time preference is a preference for *present income* as compared with *future income*; and in terms of the money economy it is a preference for *present purchasing power* as compared with *future purchasing power.*

Individuals differ considerably in respect to their preference for present rather than future income. Some can scarcely save anything at all, so great is their preference for the present. Others constantly deny themselves present enjoyments for the sake of piling up investments from which they hope to derive a future income. And they do this to such an extent that, in the opinion of their friends, they live constrained and unpleasant lives. Why these individual differences in time preference exist does not concern us very much. Like the variety of tastes and desires for present consumers' goods among different people, they are things that a preliminary study of economics takes for granted.

The same person may have, as we have already noted, a different rate of time preference under different circumstances. In youth most men are prodigal of income; in middle life they become more parsimonious. But the principal external cause of individual variations in time preference is difference in wants and provisions for wants; that is, difference in the amount of income of individuals as compared with their needs. A poor man can hardly be induced to save anything at all. If he saves he must forego some of the necessities of life, and scarcely any rate of interest could compensate for that. The rich man, on the other hand, may easily save a large portion of his income if he has not cultivated expensive habits of consumption.

The rate of individual time preference is affected by the future relationship of income and needs as well as by the present relation between these two sides of the personal budget. Man anticipates and plans for the future, and he therefore takes into account in his activities and decisions many future events. For example, if income is rising and needs in the future are expected to remain constant, the relation between income and need in the present will be affected. Under these conditions the rate of time preference tends to be high, and one may reasonably borrow to increase present consumption. On the other hand, if future needs are expected to rise and income remain constant, the individual tends to have a low rate of time preference and to save something to eke out his future income. Thus in the private economy of those who have more than enough income to buy the urgent necessities of life, there is a constant attempt to balance present expenditure and income with future expenditure and income.

---

[1] Irving Fisher, *The Rate of Interest*, p. 88. Copyright (1907), by The Macmillan Company. Reprinted by permission.

The Motives for Saving. Saving takes place because men wish to realize future objectives. In some instances these objectives are very definite. The saver wishes to spend a year in travel or to buy an expensive consumers' good that he cannot buy out of current income. Savings banks know that this is a common motive for saving and make an appeal for deposits for the purpose of buying Christmas gifts or for financing vacations. In many other cases the objective is not clearly defined. Many people save for such general objectives as security in old age, the proverbial rainy day, and the enlargement of their income and their scale of expenditure in the future; still others save to build an estate for their children or to enlarge the size of their business operations. To say that time preference exists is not, of course, to deny that men desire future income as well as present income.

Some saving would undoubtedly take place if no interest were ever to be realized from it. Needs are not necessarily distributed throughout life as income is. If interest disappeared entirely we should expect that people would still wish to provide for future needs. The parent saves while his children are young to be able to expend more for their care when, a few years later, their needs have become greater. And in many cases he would save even though no interest could be got from money deposited in the bank or invested in bonds. To do this he might hoard consumers' goods, but not to a very great extent, since commodities are usually subject to depreciation with the passage of time. Metallic money, plate, and jewels are not liable to depreciation to the same extent as are other commodities, and so we find that in backward countries, where opportunities for investment are few, saving generally takes place through the hoarding of these goods.

It is sometimes said that the rich would save even if no interest could be got from investment, because the act of saving involves little or no sacrifice on their part. There are objections to this assertion. In the first place, the accumulations of the rich are largely derived from the net income of capital or from interest itself. If interest were cut off, the savable surplus of the owners of great wealth would be diminished and saving would decline. In the second place, the statement overlooks the fact that the rich could spend a great deal more on immediate goods than they now do, though they might not spend a very much larger amount on personal gratification of the meaner sort. In Oriental countries large amounts are expended by the rulers and the rich on the maintenance of personal retainers, on the building of luxurious palaces, and on other forms of display. In the Western World the same sort of consumption takes place, but to a relatively more limited extent. The difference is doubtless due to obscure differences in the ideals and customs of the people of the two regions, but no small part of it is also due to the opportunities for the earning of interest on capital which accompany the capitalistic régime in progressive countries.

Under the social conditions prevailing in the United States, it may be true that the very rich would continue to set aside for future production the larger part of their receipts from rents and profits even though interest were considerably reduced. They would do so partly because of habit and partly because business has become for many of them an end in itself, and saving largely involuntary. We must, however, be careful not to conclude that they would continue to save as much as they now do. The habit of saving and the motives that now influence the accumulation of capital are not inborn characteristics of men and might easily be destroyed by an attempt to abolish interest.

Another form of so-called "involuntary saving," which is perhaps more important than the personal accumulations of the rich, is found in the building up of large surpluses by corporations. The net earnings of a corporation are legally at the disposal of the board of directors, who may either disburse them as dividends or reinvest them. At the present time a large fraction of all national savings results from such accumulations. Here, we are told, is a case in which the decision to save is often made by a group, the directors of the corporation, who do not own more than a small proportion of the amounts thus accumulated. But the corporate surplus often belongs to many small stockholders who have no voice in the determination of the policies of the corporation. It is said that these stockholders might not save the net earnings of the corporation if these earnings were disbursed as dividends, and that the retention of them as corporate surplus compels the small stockholders to save whether they wish it or not.

When knowledge of the fact that a surplus is being plowed back into the business is kept from all but the inner circle of the management, the argument and the conclusion are of some weight. If the stockholders do not know that the saving is taking place they cannot, of course, take any action concerning it. But there are objections to the conclusion that a lowering of the interest rate would not affect these accumulations. In the first place, if the rate of interest were much reduced, the savable surplus of the corporations would be reduced; in the second place, the directors cannot deceive the stockholders forever, and at some time the facts must be made known. Then the policy of the corporation will be changed if the stockholders do not approve it.

If we assume that the accumulation of the surplus is known to the public, then we must suppose that a majority of the stockholders are in agreement with the saving policy, and the fundamental difference between corporate savings and individual savings is not great. Probably the corporate method results in somewhat larger accumulations, since having money on hand is a different thing from having an undivided share in a corporate surplus. It is easier for the individual to spend the former. The small stockholder feels that he cannot affect the policy of the corporation, and he makes no effort to do so; but if he is greatly dissatisfied with the decision of the directors to retain the surplus, he may sell his stock. If the new owner knows the policy of the corporation, his purchase is evidence of his willingness to have the profits retained in the business.

No one saves all his income. Necessary living expenses must first be met. All in excess of this is savable surplus. Given a motive for saving, the average person then sets aside a certain sum, the size of which depends upon (1) the going rate of interest, (2) the urgency of his desire for the object for which he saves, and (3) the rate at which the marginal utility of the income remaining for present consumption rises as more and more is diverted to future uses. At some point another $100 added to savings makes possible a future objective that just balances in desirability with the desirability of the present goods the $100 would buy. At this point the saver ceases to add to savings and consumes the remainder of his income, and at this point his rate of time preference is equal to the rate of interest at which he can invest his savings.

The Supply of Savings. The total supply of savings, — that is, the supply of current income offered for financing the building of new capital, — is made up of the supplies of numerous individuals. As we have seen, the amount any individual will offer varies with the rate of interest; and the amount the entire body of savers will bring to the loan market also varies with the interest they can get, just as the total supply of a commodity varies with the price that the consumers are willing to pay. Like the supply of any commodity, the supply of savings may be affected by conditions other than the interest rate (price in the case of a commodity). Changes in the attitude of parents toward their children, changes in the standards and modes of consumption, changes in the distribution of the national income, all have some effect on the price at which different parts of the total annual savings will be brought to the loan market. At present we are concerned with the relation between the rate of interest and the volume of savings.

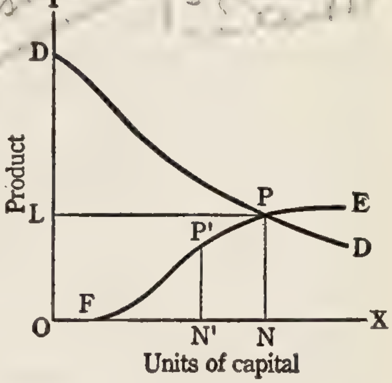

FIG. 37

Some part of the supply of savings would be forthcoming even if no interest were paid, and some if a substantially lower rate than the one prevailing were offered. But the going rate of interest is necessary to bring the supply of savings up to its present volume. The supply of savings may be illustrated by the curve *FE* in Fig. 37. If only a very small amount of savings were demanded, it could be obtained at zero interest. If *ON'* units are to be supplied, *N'P'* interest is necessary; but if *ON* units are required, *NP* interest is necessary. And the rate which would correspond in percentage to *NP* is the marginal cost of saving. The area *OLPN* is the extra amount of future goods that must be paid in an exchange of present goods for future goods. For since all savers are paid at the same rate, not only the marginal increments but also all other portions of the annual total of savings will receive *NP*.

It must not be supposed that the curve *FE* represents the interest rates at which different people or groups of people can be induced to save successive amounts: it is supposed to show the cost of calling forth different amounts of saving from the entire community. One person might save a small amount at zero interest, a larger amount at $N'P'$, and a still greater amount at NP; for others the minimum effective rate might be $N'P'$ and for still others NP.

The marginal cost of saving is not a fixed, or constant, amount. Changes in the income of the mass of savers will affect it. If spending habits remain constant and if income increases, the rate which would call forth a given volume of savings will fall. If, on the contrary, income diminishes, the rate will rise. Wars and other agencies that destroy capital or reduce income tend to raise the marginal-supply price of saving. Changes in the habits of consumption of a people will also affect the time preference of a majority, a tendency to easier living and spending being accompanied by rising rates and a reverse tendency by falling rates. Thrift can be taught, and there is no doubt that an intelligently conducted campaign for the saving habit would bring results. Unthrift may also be taught, of course, and every seller of luxuries who by advertising and other means persuades people to buy his wares is engaged in teaching it.

The Demand for Savings. The demand for savings arises from two sources. In the first place, many people want loans to buy consumers' goods. Their demand gives rise to consumption loans. In the second place, entrepreneurs want present purchasing power to invest in capital goods. The latter is by far the more important demand and it is the one that is predominant in determining the normal rate of interest. The demand of the entrepreneur for loans is based on the productivity of capital. To return to an earlier example, if the entrepreneur can buy for $1000 a machine that will, in the course of its life, earn enough to repay its original cost and at the same time yield a surplus of $60 a year, then he can afford to offer 6 per cent for a loan of $1000 in order to make the purchase. That is his maximum demand price for a loan that enables him to buy the machine.

If we should take a limited view of the loan market, we might conclude that the explanation given in the preceding paragraph was a satisfactory account of the demand for savings. But this explanation assumes that the machine which will earn replacement and $60 a year can be bought for $1000. Now the value of any producers' good depends on (1) the value of the products, (2) the rate at which the value of the products expressed in money is discounted, and (3) the distribution of the products in time. The value of the products of any capital good depends, as we have seen, on the demand for them and on the amount brought to market. But the product of the machine is not a separate product, and the value that the entrepreneur will impute to the machine depends on the marginal productivity of the machine. If any kind of capital good becomes more plentiful, its marginal productivity will decline, and the entrepreneur will pay less for it.

The money cost of any item of capital goods, say a machine, arises from the fact that there are rival demands for the factors used in making it. Because of these rival demands the owners of the factors can exact a payment for their use in making capital goods. Even if interest did not exist, and if capital goods earned only enough to replace themselves, the law of opportunity costs would operate. But when the investment in the various sorts of capital goods must earn interest, it is evident that the earnings of these goods must exceed their costs. In Chapters XXIV and XXVI, it was explained that the marginal productivity of labor and of land determined their earnings. The same principle governs the earnings of concrete capital goods.

But at this point another fact must be taken into account. Although capital goods are the products of the primary factors of production, their productivity is not the same as the productivity of these factors. The capitalistic method, as we have seen, increases the productivity of the primary factors, and the business demand for saving arises from this fact. If savers can be found who are willing to wait for the products of the machine and allow labor to enter into the roundabout method, more products or more valuable products can be created than if all the labor is used directly. And as long as there are methods of attack on the environment which give larger results in terms of value than direct methods do, entrepreneurs will demand capital goods in order that they may take advantage of these methods; and as long as *new* methods can be discovered, the demand for additional capital will continue.

As machines and other capital goods are multiplied the marginal productivity of each machine becomes less and less, following the universal law of diminishing returns. To put it in another way, the labor and land which go into the machine become less and less important as compared with the labor and land used directly as the former are increased in amount. Not all capitalistic processes are equally productive. Some excel the direct method to an almost unbelievable extent; others add but little. In any given stage of the development of technique the most productive are first selected for exploitation. But a point is finally reached where the multiplication of capital goods of the more productive sort has gone so far that the amount by which the machine method excels the direct method is just sufficient to offset the interest charge which the savers require to induce them to wait for their income. At the same time that these machines are being multiplied, other less productive machines and capitalistic processes will be brought into use in industry. Many of these will not be as productive as others. There will be all degrees of gain over the direct method. We may call these latter processes inferior, just as we call the less productive areas of land inferior. The decline in the additional product from these increasingly inferior machines marches along with the decline in the productivity of the additional machines of the superior sorts. Both lines of development finally reach the point where the superiority of the capitalistic process is just sufficient to offset the required payment for waiting.

In terms of Fig. 37 the demand of all society for saving is illustrated by the curve *DD*. As saving increases, the purchasing power which is turned over to the entrepreneur is invested in capital goods of less and less productivity, indicated by the slope of *DD*, until the marginal unit at *NP* just yields a net return equal to the interest that must be paid to call it forth. The curve *DD* in this figure does not represent the gross productivity of capital goods, but only the excess productivity of the capitalistic method over the direct method. The marginal unit of capital goods must yield, in addition to its cost of replacement, a surplus measured by *NP*.

At this point new saving ceases, for further accumulation would not pay the premium demanded by savers in the exchange of present goods for future goods. Whatever quantity of capital goods has been brought into existence at this point is the amount that labor and entrepreneurship will have to work with. It is also the amount that will be applied on the land. The marginal cost of saving limits the supply of capital goods. And that supply in turn determines the marginal productivity of capital goods. It partly determines the wages of labor, the earnings of entrepreneurship, and the rent of land; but only partly, because the productivity of these factors is also affected by the supply of them. The earnings of labor, land, capital goods, and entrepreneurship are mutually determined; that is, wages depend on the marginal productivity of labor, but how much a given number of laborers can produce depends on how much land, capital goods, and entrepreneurship are combined with them. The rent of land is also dependent on the supplies of labor, capital, and entrepreneurship which are combined with it. Given the state of technique in production, the productivity of each factor depends on its supply and on the supplies of the other factors.

We now return to the question What determines the price of the machine purchased by the entrepreneur? If we are concerned with a single machine or type of capital, the answer is that the machine must sell for its cost of production, including wages of labor, rent for land, the earnings of entrepreneurship, and the returns to other capital; but if we substitute for a single type of capital *all capital,* the answer is that the cost of capital goods is simply the value of labor and entrepreneurship working on land without the assistance of capital. This is not entrepreneurs' cost; for entrepreneurs do not deal with capital in the mass, but only with particular units of capital goods. But from the social point of view the cost of production of capital is as has been stated. This explanation seems at first to be incorrect, because we know that without the assistance of capital goods labor could produce very little, and the cost of capital would therefore seem to be much too small. But as labor is withdrawn from direct production the marginal productivity of the remaining labor rises, and the productivity of the labor which is embodied in capital goods falls as more and more of it is put into them and the capital goods in which it is invested are multiplied. A balance is reached when a further subtraction of labor from the more direct methods would destroy a marginal product that was greater than the marginal product added by capital discounted to the present.

The Capitalization of Incomes. When the income from any durable agent of production is reduced to present worth, that income is said to be capitalized. A tract of land, for example, the annual rent of which is $100 is worth $2000 when the interest rate is at 5 per cent. This is the capitalized value of the annual rent forever. It would be worth more than that if the rate of interest were lower, and less if the rate were higher. The income from capital goods may also be capitalized; but since these goods are seldom as durable as land, and since they must all be produced, the phenomenon of capitalization does not play as important a rôle in determining their values as it does in the case of land. If the capitalized value of the future earnings of machines seems unlikely to amount to the price asked for them, the entrepreneur will not purchase. If the capitalized value of the machines is in excess of the price, the demand will cause the price to rise; but the rise in price will be answered by an increase in supply; the added supply of machines will reduce their marginal productivity, earnings will decline, and the capitalized value of the machines will approach their cost of production. The rate of interest thus acts as a selective factor in determining which machines or other capital goods shall be produced.

When capital has once been installed, it will continue in use as long as it has any marginal productivity. This statement needs some explanation, because we know that capital goods are often discarded before they have ceased to be useful, and better equipment is put in their places. But in these instances the capital is really not productive. It is either so far worn out or so antiquated that it detracts from the productivity of the labor and land with which it is used, and cannot be said to earn anything. Old capital goods, before their abandonment, may yield a return that is much below the gross productivity required to replace the original investment in them and pay the going rate of interest. In short, they yield a return that, like the rent of land, is not related to their cost of production. To illustrate this statement we may suppose that an old piece of capital yields $200 a year, which exceeds the replacement charge by only half the amount necessary to pay the going rate of interest on the undepreciated portion of the original investment. Its value is now determined by its diminished earnings, capitalized at the going rate of interest for the period of its life, and not by its cost. In the same way, the value of capital goods which have been installed by incompetent entrepreneurs does not depend on cost; neither does the investment of savings in these goods yield the going rate of interest. Strictly speaking, only new, well-planned investments in capital goods yield interest; old and badly planned investments yield a return that resembles rent. Sometimes this rent affords a net return in excess of the going rate of interest, since it is conceivable that the capital goods might yield a return in excess of replacement cost and the going rate of interest. But this condition cannot persist very long. If a machine costing $1000 yields something more than replacement and the going interest charge, its supply will be increased, its marginal productivity will decline, and its earnings will fall to the amount which when capitalized is only equal to its cost.

Types of Interest. We have used the term "interest" to mean the payment necessary to equalize the values of present and future goods. The typical example is found in the case of a loan to a well-established government. Here the lender turns over to the borrower a certain sum of money and is assured by his contract that at stated intervals he will receive interest, and that at the expiration of the term of the loan his original investment will be returned to him. There is a minimum of risk attaching to the loan, and the conditions of the contract are clear and explicit.

The investment of the entrepreneur in capital goods does not involve a contract between a borrower and a lender, yet the investor expects and ordinarily receives payment for waiting. This is called *implicit* interest to distinguish it from the payments made on a loan contract, which are called *explicit* interest. The investor in stocks may receive two types of return — implicit interest and pure profits. The former is determined according to the demand for waiting and the supply of waiting; the latter, according to the principles governing pure profits. In numerous cases the common stockholder receives nothing at all. But the same may be true of those who buy bonds or lend on other security. All investments involve a risk that neither interest nor principal will be repaid. To cover this risk the investor requires an addition to the riskless, or pure, rate of interest. When government bonds selling at par yield only 4 per cent the bonds of corporations whose future earnings are insecure often yield 6 or 7 per cent. The addition of 2 or 3 per cent to the riskless rate is to cover the losses which the investor believes will occur, on the average, in respect to the risky securities. If these estimates of risk were accurate, then a large sum of money invested in government bonds and the same amount invested in poorer bonds would in the long run bring in about the same amount of interest. But the average investor cannot distribute his holdings of the poorer bonds widely enough to obtain the benefit of the law of averages. Hence if he is lucky he may receive more interest than if he had bought government bonds; if he is not lucky he may receive less.

Some investors cannot afford to take large risks. Their savings are small and the consequences of loss are too serious. Other investors can afford to take greater chances. The supplies of savings are not homogeneous. It follows that the interest rates yielded by very secure bonds and by more risky securities may differ by more than the true insurance premium.

The Loan Market. In the preceding sections of this chapter we have intentionally refrained from introducing any discussion of the manner in which savings find their way into the hands of entrepreneurs, and we have not dealt in terms of market rates of interest. We have been concerned almost altogether with the problem of normal, or long-time, tendencies. When an individual saves he refrains from spending all his income on consumers' goods and invests the remainder in some form of capital. He may do this by actually buying the capital for use in his own business. From the gross receipts of production the manufacturer must pay all expenses of production. If a surplus remains he may either spend it all for consumers' goods or he may buy more equipment or more land. If he buys land he merely displaces some other owner; if he buys more equipment he causes more of these capital goods to be produced.

In many instances the saver cannot invest directly by buying capital goods: he must rely on the loan market for his investments. Having a savable surplus, he may take life insurance, buy bonds, or merely leave the money with a bank. In any case purchasing power is now available for investment in new capital. But the saver does not make any calculation as to the probable future return from capital.

The entrepreneur takes the responsibility for choosing the capital goods that shall be bought with the purchasing power which the saver does not wish to spend on consumption. He makes a contract with the saver to pay him a stipulated rate of interest and to return the principal of the loan at a definite date in the future. Sometimes the investor buys stocks and not bonds. But the investor in stocks does not determine the kind of physical good that shall be added to the equipment of the corporation. The managers of the corporation, acting as his representatives, do that.

When one saves he causes purchasing power to be used to produce capital goods rather than consumers' goods; or, more accurately, he brings about the production of future income (products and services) rather than present income. The loan is usually made in purchasing power; that is, in money or, more frequently, in deposit money. And the supply of savings that is seeking investment at any time is the supply of purchasing power which is offered for investment in capital goods and for consumption loans.

The agencies through which savings are directed to the loan market are numerous. Banks, trust companies, mortgage companies, bond houses, investment trusts, and life-insurance companies are in the market offering the accumulations of the savers to entrepreneurs and consumers who wish to borrow. The supply of these savings is met by the demands of the entrepreneurs and consumers who wish to make loans. And at any time the rate of interest is determined by the interplay of the forces of demand and supply. The demand of the borrowers is momentarily determined by the condition of trade, by speculative activity in the community, and by all the other forces that bring about opportunities for buying and selling at a profit.

The demands of the various borrowers are not homogeneous. Merchants and manufacturers make use of a great amount of short-time credit to finance purchases of stocks and of raw materials. They also desire long-time credit to buy permanent equipment. For the former they turn to the banks; for the latter they turn to private investors and to trust companies and other financial agencies which make a business of acting in the rôle of go-between for savers and entrepreneurs. There is a more or less clear distinction between the markets for short-time and long-time loans. Moreover, investments for a longer period of time are distinguished as to security (first-mortgage bonds versus second-mortgage bonds etc.) and as to the length of time for which the loan contract runs. There is also a distinction between the market for loans of large corporations of established reputation and the loans of smaller concerns which are equally sound but which are not so well known. Nor are these more than samples of the various types of demand for loans in the modern loan market. We find that instead of one kind of demand for the purchasing power of the saver, there are many. Furthermore, as has been said, savers are not indifferent to the kind of investment they make. So instead of a single loan market, there are many, and in each of them there is at any time a supply of savings seeking borrowers and a demand for savings seeking lenders. The market rates of interest are determined, like other market prices, by the action of demand and supply.

Interest Rates and the Supply of Money. One of the early explanations of interest ascribed fluctuations in the rate to variations in the supply of money. Since loans are usually contracted in terms of money, it was perhaps natural that men first attributed these fluctuations to the increase or decrease in its supply. In more subtle forms the explanation persists at the present time.

One special argument for the money theory of interest attaches particular importance to the supply of bank credit. It is well known that when the reserve of a bank increases, its loans may also increase if the banker sees fit to expand them. And the loans may be increased more rapidly than the reserve, since only a fraction of the deposits resulting from loans needs to be covered by reserve. Here, it is said, is a case where the supply of loanable funds depends on the supply of money or, at least, on the supply of substitute money — bank credit. If we examine this statement we must agree, in the first place, with the assertion that an augmentation of reserves will probably result in larger loans at lower contract rates of interest. The banker having the opportunity to do so will generally seek to expand his earning assets, loans, and discounts. The market rate of interest on short-time loans will therefore be affected by an increase of bank reserves. But this effect will not persist. In the first place, the borrowers who are induced to make loans by the reduction of rates will now compete for the physical factors of production and force up their prices. This in turn will necessitate larger loans to finance the same physical volume of trade, and in the end the rise in prices will compensate for and offset the added reserves. In fact, the rising-price level usually raises the rate of interest on bank loans. Entrepreneurs gain during periods of rising-price levels and are therefore eager to expand their businesses. To do this they need larger loans; and as long as their profits continue to expand, all but the wariest of them will continue to borrow. The augmented demand resulting from rising prices, which result in turn from an increase in the supply of money, drives the rate on bank loans temporarily higher than it would have been had no increase in the quantity of money taken place.

But the rise in the price level affects savings. It does so because the increase in bank credit arms the entrepreneurs with more purchasing power and at the same time, through the effect it has upon prices, deprives those who receive fixed incomes of a part of the purchasing power which they formerly enjoyed. The entrepreneurs use the increase in purchasing power which comes to them to increase capital goods at the same time that the power of the fixed-income groups to buy consumers' goods is reduced, and the result is an addition to the capital of the community somewhat in excess of the addition that would have been made if the level of prices had remained constant. But the effect of this process upon the accumulation of capital in the community must not be overstated. Some of the receivers of fixed incomes, notably the well-to-do bondholders and landlords and the better-paid professional classes, are savers. When the rising-price level reduces the purchasing power of the money incomes of these groups it is probable that they reduce both their savings and their consumption. The amount by which they reduce their savings is an obvious offset to the accumulations by the entrepreneurs to whom the banks have extended credit.

If there is a net addition to the capital of the country the marginal productivity of capital generally will tend to decline and the demand price for loans will also tend to fall, for a time, at least, because entrepreneurs will be unable to earn enough on the added increments of capital to enable them to pay the old rates of interest. The enforced savings thus produce a decline in the market rates of interest. The decline in these rates will set in motion other forces, however, which will tend to bring about a rise. The incomes of many savers will be reduced and some may refuse to save and turn to consumption because of the decline in rates.

But in the long run interest rates must seek the point where the demand for savings, which is determined by the net marginal productivity of capital, and the supply of capital, which is determined by the marginal cost of savings, are in stable equilibrium. The values of fixed-capital goods will, of course, adjust themselves to changes in the rate of interest. The bank rate, although its fluctuations are governed by many other forces, must in the long run agree with the rate of interest which will equate the demand for and the supply of savings. If it is temporarily below this rate of interest, loans will be expanded, the price level will rise, and the bank rate will also rise. If the bank rate is above this rate loans will be contracted, prices will fall, and the bank rate will fall. Thus the bank rate is only temporarily controlled by the volume of money and the reserves of the banks. In the last analysis, the bank rate must seek the level of the rate of interest that capital can earn in the productive process.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Interest was once explained and justified on the ground that the lender had sacrificed the opportunity to use in his own business the funds he lent. Is this a valid explanation?

2. An increase in the productivity of capital goods, say by an invention, would increase the demand for saving. Why? Would it also increase the rate of interest for a time? Would it cause the normal rate of interest to rise or to fall?

3. The banks of a certain city have for some time made use of advertisements to induce people to deposit on savings accounts. Do you think this type of advertisement is likely to prove effective? If it does, will it affect the sales of merchandise in the city?

4. Suppose the productivity of a large area of land were increased by the building of a railroad. How would the rent of the land be affected? Would interest rates in the community rise as a result of this increase in productivity?

5. How can a pawnshop proprietor exact several times the contract rate of interest that is charged by a bank? Does he actually receive higher interest?

6. "People do not save to obtain interest but to accomplish certain definite objectives — to build houses, to educate children, etc." Is this statement correct? Is it partly correct?

7. In new countries the rate of interest is usually higher than in old countries. How do you explain this? Is it because capital is more productive in new countries? because the supply of savings is less? because time preference is higher?

8. Is a demand for dwelling houses a demand for savings? a demand for short-time consumers' goods, such as clothing? Explain.

9. Suppose that the rate of interest should greatly increase. How would that affect the value of land? of bonds? of automobiles?

10. Great wars are usually accompanied by rising interest rates. This is due in part to greater risk in connection with all investments. Is this the only explanation?

11. How would the rate of time preference be affected by (1) a steady growth of the national income? (2) extravagance in consumption? (3) old-age pensions paid by the government?

12. Suppose the interest rate should fall to 2 per cent. How would this affect the rentals that tenants pay for houses?

CHAPTER XXVIII

PROFITS

Business Profits and Economic Profits. The term "profits," as it is sometimes employed in accountancy and generally in business, includes a number of distributive shares. It is a name given to the net earnings of the proprietorship after all necessary payments connected with the business have been made to outsiders and after expenses for depreciation have been deducted. The following statement is illustrative of the derivation of net profits for a small manufacturer:

Gross earnings		$155,000
Deductions from gross earnings		
1. Cost of materials	$30,000	
2. Labor	40,000	
3. Other manufacturing expense	5,000	
4. Cost of selling	10,000	
5. Depreciation on plant	8,000	
6. Taxes	6,000	
7. Interest	11,000	110,000
Net income or business profits		$ 45,000

The remainder from gross earnings after all deductions is clearly a composite of several economic distributive shares. Since the manufacturer owns the plant, it includes interest on capital and rent for the site on which the plant stands; it also includes a portion of the wages of the owner of the business. This composite return, the net gains of the proprietorship, we shall call *business profits*.

By *economic profits*, or *pure profits*, we mean the gains of the entrepreneur over and above all explicit expenses of conducting the business and all imputed costs, which include depreciation of the plant, the competitive interest on the proprietors' investment, the competitive rent on their land, and the wages they could earn as salaried managers. Recast to conform to our definition of pure profit, this statement stands as follows:

Gross earnings		$155,000
A. Deductions for payments to others		
1. Cost of materials: wages, rent, interest, and possibly profits	$30,000	
2. Labor: all wages	40,000	
3. Other expenses: wages, rent, interest, and possibly profits	5,000	
4. Cost of selling: wages, rent, interest, and possibly profits	10,000	
5. Depreciation: wages, rent, interest, and possibly profits	8,000	
6. Taxes: wages, rent, interest, and possibly profits	6,000	
7. Interest: chiefly pure interest, plus some wages and possibly profits	11,000	
Total	$110,000	
B. Deductions for imputed costs		
1. Wages	$10,000	
2. Interest	25,000	
3. Rent	5,000	
Total	$40,000	
Grand total of all deductions		$150,000
C. Pure profits		$5,000

In this account pure profits are much smaller than business profits in the preceding account. This difference is due to the deduction of imputed wages, interest, and rent. But the difference is more fundamental than mere procedure. Business profits do not vary entirely with the efficiency, competitive position, or luck of the concern, whereas pure profits do. Two businesses might show different amounts of business profits simply because the investment of the proprietors was larger in one case than in the other; but when allowance has been made for the imputed distributive shares, no discrepancy can arise from the fact that one concern owns its plant and equipment free from all encumbrance, whereas the other has borrowed a part of the money necessary to purchase these properties.

The existence of a pure profit year after year is almost conclusive proof that the concern receiving it either has able entrepreneurs at its head or holds a position which cannot be attacked by competitors.

Pure Profits as a Share in Distribution. The three distributive shares,—wages, rent, and interest,—accrue to the owners of the three factors of production to which they are ascribed because of the productive contribution of these factors. If the fourth factor, entrepreneurship, is a productive agent, it must likewise share in the product of industry, and we must discover for it a marginal contribution to that product. This factor is, however, not entirely comparable to the other three. In Chapter IV the functions of entrepreneurship were briefly described. It was there explained that these functions are always exercised by human beings, but that the work of those who perform these functions is not the same as the work of even the most highly paid routine worker. Nor are the qualifications identical for the two kinds of work.

Alfred Marshall has characterized the work of a successful independent manager as that of informing himself concerning the markets, materials, and productive processes of his business; of forecasting demands and costs; of meeting new situations with quick decisions and prompt action; and of selecting good subordinates and inspiring them with the spirit of coöperation.[1] The Austrian economist Wieser has also stressed leadership as the distinctive characteristic of the independent manager.[2] Other economists, emphasizing the work of organization, have found the distinctive productive function of the entrepreneur in the minimizing of losses by careful combination of the agents of production and by the direction of these agents to their most productive uses.

This characterization differentiates the function of the entrepreneur from that of the labor of the ordinary routine sorts, but obviously it does not establish a clear *distinction*

---

[1] *Principles of Economics* (Eighth Edition), pp. 297–298.

[2] *Social Economics*, p. 324.

of kind between entrepreneurship and the highest sort of direction and control by salaried employees. Many plant managers, railroad executives, and managers of large mercantile concerns must have the characteristics and do the sort of work that has been described; yet they receive wages, or salaries, and cannot, so far as the theory of distribution is concerned, be said to exercise a productive function different in kind from that of ordinary labor. If we read the biographies of successful entrepreneurs we find that many of them began as salaried workers in not very responsible positions, were later promoted to managerial or executive jobs, and finally became owners and responsible heads of their own businesses. Nowhere is there a distinct break between the work they did in the subordinate position and the work they were called upon to do in the higher position. The differences are all of degree and not of kind.

Land, labor, and capital are distinguished from each other by (1) the differences in the parts they play in production and by (2) differences in the conditions governing their supplies. The functions of land and labor are especially unlike: the former is the passive agent in production, whereas labor is the active agent. Land furnishes the materials of production; labor moves them about — reforms and recombines them. The supply of land is dependent on unalterable natural conditions, discovery and changes in technique excepted; labor in a technical sense is a wasting, replaceable agent. The supply of land is unrelated to cost of production, whereas the supply of labor depends, in part at least, upon its costs of production, both nominal and real. Capital is, from the physical point of view, a product of the other two factors, but its supply is limited by time preference, which is a sufficient distinction to warrant its being set up as a separate factor in production. Entrepreneurship cannot be so clearly differentiated. Its supply is controlled by the same forces that control the supply of labor in general. There are differences between the conditions that make for a large or a small supply of entrepreneurs and other laborers, but the conditions affecting the supply of the former and the supply of professional workers and salaried managers occupying important positions are substantially the same. Neither is there a striking difference between the work of the self-employed business man, whom we must call an "entrepreneur" or else abandon the term, and the work of the most responsible managers.

It is sometimes said that the real difference between the position of the self-employed business man and the salaried manager is the fact that the former must take sole responsibility for the entire business, whereas the latter is responsible only for the proper conduct of the office he holds, and that if he fails the burden of his failure will fall on the shoulders of the real entrepreneur. But even this distinction disappears almost entirely when the corporate form of organization supersedes the individual form. If the managers and the board of directors of the corporation fail, it is the stockholders or both the stockholders and the bondholders who suffer and not someone higher up in the management of the business. The directors suffer because they have been owners of the securities of the corporation and because they will be less likely in the future to secure positions of trust and responsibility, but both sorts of loss might fall on a salaried manager. There is no escape from the conclusion that the entrepreneur is a laborer. *But there is a difference of degree.* The common laborer is held responsible only for the simplest sort of routine work; the skilled worker is assumed to know how to do more difficult tasks; the salaried manager is expected to know routine and to be able to judge conditions that cannot be described and provided for in advance; and the entrepreneur is required, if he is to be successful, to meet still more uncertain conditions outside the business and to control men and materials within the business under situations that are not easily dealt with by routine methods. The difference in the work of the entrepreneur and of the highest type of salaried official may be very great or it may be slight.

Profits from the Earnings of Rare Natural Ability. In business, as in the professions, men of exceptional ability make very large incomes. The lawyer whose ability to analyze legal problems or to plead cases excels the ability of all but a few of his competitors receives fees greatly in excess of the average practitioners. So too the entrepreneur whose business ability exceeds that of his typical competitors by a wide margin will receive a labor income greatly in excess of that made by a majority of the trade in which he is located. His marginal productivity is very great because under his direction the other three factors of production are more productive than they are under the management of men of ordinary ability. He is responsible for these earnings, and they would disappear if he were withdrawn from the combination and a manager of inferior ability put in his place.

It may be asked, however, how this sort of profits differs from wages. For the most part it differs not at all. Yet there is one important distinction between the earnings of the self-employed business man and the earnings of the salaried manager. When the latter is hired, the owner of the business agrees to pay a stipulated salary for a given length of time, regardless of the success or the lack of success of the management. The salary will not be continued, of course, if the manager is unsuccessful; but there is always the risk that the manager will not be successful, and in that case the immediate loss will have to be borne by the owners of the business.

The earnings of executives who possess unusual capacity for leadership in business will always be high, either when they are employees or when they are self-employed. Usually they are not found in the ranks of the employee class. Not only do their energy and initiative drive them into independent undertakings, but the recognition of their abilities by others enables them to obtain capital so easily that they are seldom allowed to remain in salaried positions. Some retain the status of employee, but at the same time exercise dominant control of the organizations for which they work; still others ability [illegible]

retain salaried positions and at the same time conduct independent businesses in other lines.

But the profits of business enterprise do not arise entirely, under present-day conditions, from the earnings of exceptional ability to manage business under difficult conditions. Stockholders in corporations often receive from their investments a return greatly in excess of the going rate of interest, and the same is true of owners of small businesses who are not conspicuously more adept in management than thousands of other men who receive only the going rate of return on their investments in capital. For the most part these extraordinary gains arise from two sources: from the unavoidable risks and uncertainties of business and from interferences with free competition. We shall turn, first, to the effect of risk and uncertainty on the returns to the factors of production.

Uncertainty in Economic Life. In a changeful, dynamic society uncertainty is always present. Every person who engages in any economic activity must encounter situations in which the effects of action, of decisions, and of policies cannot be definitely forecast. The laborer takes the chance of bodily injury, of unemployment and abnormally low wages because of an oversupplied labor market, and of adverse changes in the real value of his money wage. The skilled laborer encounters uncertainty in respect to all these, plus the further chance that the trade which he has learned at a considerable money expense and at the sacrifice of time and energy will be rendered obsolete by an invention or a change in the demand of consumers. Labor in the salaried class or in the professional and managerial groups runs less risk of temporary unemployment and of bodily injury than do factory hands and ordinary salespeople; but in these higher occupations there are also uncertainties of demand for the services of the workers and chances that the type of work for which they have prepared themselves will become obsolete.

The nonhuman factors of production are also subject to uncertainty. Owners of land cannot completely foresee the growth of business in the city where the land is situated, the direction of spread of population and of business-building, the amount of taxation, and many other conditions that affect the rental of their property. Capital invested in any productive enterprise must be so invested in the face of numerous uncertainties. Machinery and other capital equipment may be superseded by improved equipment. The owners may have located the plant in a good territory at the time the plant was built, only to find later that a new plant in a better location has deprived them of their markets. To conclude: all types of economic activity and all the factors of production are subject to uncertainty.

The term "risk" is commonly used to mean the chance of an *adverse uncertainty*. One would not speak of the "risk" that his house would not burn, or that the growth of business in the neighborhood of his land would cause his rentals to increase unexpectedly. Hence we have used the term "uncertainty" to indicate that the unforeseen future events may turn out either more favorably or less favorably than had been anticipated. There is no positive gain to offset bodily injury, — one of the risks of labor, — or to offset loss of capital by fire, windstorm, and the like. These are risks from physical injury. But purely economic uncertainties are of the nature of variations about normal price. The laborer may suffer unemployment, or he may chance to obtain work under conditions where the wage is adjusted on the assumption of a certain amount of unemployment, and yet not be subjected to loss of working time. The owner of capital may find that the demand for its products increases unexpectedly and that his gains far exceed the marginal cost of saving. Landowners receive windfalls from various sources, all of which increase the rental of the land by increasing the demand for it. Purely economic uncertainty gives rise, then, to both unpredictable gains and unpredictable losses.

The fact that uncertainty surrounds the utilization of the factors of production does not mean that each and every owner must take chances of loss or gain in employing them. The owner of buildings may shift the risk of loss by fire to an insurance company; the saver may transfer some of the risk of loss to others by lending to them on mortgage security greater than the amount of the loan. Throughout the economic structure of modern society we find that the taking of risk — the assumption of responsibility for meeting uncertainty — is shifted from groups owning certain factors to other groups who own either different factors or similar factors. But the people who chiefly relieve others of uncertainty are the owners of property — land and capital. Laborers cannot easily guarantee others against loss. Human beings are not vendible assets, and no mortgage can be placed on them to secure other laborers or the owners of land and capital. Examples of the transfer of uncertainty-bearing from one group to another are numerous. When the laborer receives a stipulated wage, he transfers to his employers the risk that the goods he produces will not sell for anticipated prices. The investor in bonds and in mortgage notes shifts the uncertainty that pertains to the properties in which his money is invested to the stockholder and to the owner of the mortgaged property. But there is no complete shifting. The employer may become bankrupt and fail to pay the stipulated wage; the corporation may fail and the security for its bonds prove inadequate; and the property by which the mortgage note is secured may turn out to have less money value than the amount of the loan.

The Entrepreneur as Risk-Taker. It usually happens in a small business that the responsible head has made a considerable investment of his own capital. This is but natural, since neither laborers nor investors will intrust their services or their property to one who has no assets with which to guarantee, at least partly, the payment of wages and interest. Final authority and the power to guarantee usually go together. The control of a business is not usually intrusted to men unless they can be held responsible for their decisions, and the final test of responsibility for policies and minor decisions is the ability to make good such losses as arise from errors in judgment with valuable property.[1] Thus it comes about that the entrepreneur is both the final authority and the principal risk-taker in the average small business. Naturally this arrangement is satisfactory to the entrepreneur as well as to the investor and the laborer. If a man is asked to pledge his property for the losses that may occur in the conduct of business, he wishes to have the final decision on questions pertaining to the conduct of the business.

The development of the giant corporation has changed somewhat the relations between risk-taking and final authority for business decisions. As we know, the stock of these concerns is often held by thousands of people who own only a few shares each. But a majority of these widely scattered owners of stock never participate directly in the determination of business policies. A small inner group who are large stockholders, or who can count on voting by proxy large blocks of stock, or who have secured control in some other manner run the affairs of the corporation. Here we find a partial separation of final authority from risk-taking.

Risk-Taking and Profits. If one person relieves others of the unavoidable risks of business, he must bear the losses that arise from adverse events; he is also the recipient of gains arising from favorable events. If an undertaking turns out better than had been anticipated by a majority of those who have opportunities to invest in it or similar enterprises, pure profits are made; if the outcome is less favorable than had been anticipated, the result is a net loss. Now it might be supposed that after a certain type of business had been operating for a generation, the chances of loss and gain would become known, just as it becomes known about how great the probability is that a certain number of houses in each city will burn annually. Something of the sort does take place. Old lines of business are less uncertain than new ventures, which are proverbially speculative. But purely physical phenomena, such as fires and windstorms, against which insurance can be carried, are the results of relatively stable forces. In purely economic affairs there is not the same stability. Even well-established lines of business are subject to the effects of changing demand, inventions, shifts in the location of population, and many other variations that affect costs of production and price. Nothing like the general experience of the fire-insurance companies emerges. Moreover, no insurance scheme will enable one to predict whether a given building will burn. Only the general average liability is ascertainable, and even that is not accurately predictable. Whether a given business enterprise will encounter what might be called "average conditions" is always a matter of uncertainty.

---

[1] There are of course exceptions — trustees, executors of wills, guardians of minors, and the like.

Pure profits arise in part from the good luck of certain enterprises. These are the concerns that encounter less than the average adverse conditions. And the profits that come to the proprietors are purely *fortuitous gains.* But since the laborers and the bondholders have been guaranteed wages and interest, the stockholders retain the larger portion of these gains. The investors on mortgage security may likewise secure a part of them, because all contractual interest rates are affected by the risk that the principal will be lost if the mortgagor fails in business. To put the matter concretely, if the investors have lent money to the concern at 6 per cent when government bonds paid only 4 per cent, the extra 2 per cent is chiefly a premium for risk-taking. If the investors find that after half a century no loss has occurred, they may count a part of their receipts from contractual interest as pure profits.

In dealing with conditions of uncertainty some men have far greater ability than have others; some have far greater opportunities to discover good investments than have others; and some have better opportunities for spreading their investments widely, so that, like an insurance company, they will receive the benefit of numerous chances. If the individual has exceptional ability to estimate uncertainty, the extra return on his investments must be attributed to his personal abilities and is therefore a kind of wage. It is substantially the same as the pure profits that arise from exceptional ability as an entrepreneur. The gains which an individual receives because he has better opportunities to know the relevant facts about investments than a majority are clearly gains which arise because of the impediments to competition. The same conclusion applies to gains that come from the possession of a large amount of wealth which enables the owner to spread his investments.

The Real Cost of Risk-Taking. If the assumption of the greater part of the risks of business were regarded as a hardship by the entrepreneurs, then a special compensation would need to be forthcoming before men would assume these risks. It is often argued that men prefer certainty to uncertainty; that, other things being equal, one would rather have a good investment that paid 5 per cent than a risky one that promised either 10 per cent or nothing. But this generalization is not accurate. Investors are not all of the same temperament. Some do actually prefer the chance of great gain, even though it is a very remote one, to an investment that promises a smaller return as a maximum but a larger probable one. Others can scarcely be induced to take any chance at all. They too overlook the actuarial probability. Investors of the cautious type are not usually found among the entrepreneur class unless the inheritance of property forces them into the position. In business they seek salaried positions, and for their investments they demand conservative securities. Entrepreneurs are recruited from those who are willing to take risks.

Whether payment is necessary to induce the bulk of the entrepreneurs to enter business cannot be definitely determined. There seems to be little evidence that anything in addition to the chance that they *may* make large fortunes is necessary. And this conclusion is given more weight by the fact that the independence enjoyed by the entrepreneur is to many men a very great satisfaction. They wish to give orders, not to receive them, and they wish to manage their own investments. But independence and management are attended by risk, and with control goes responsibility.

Natural Limitations on Competition. The theories of value and distribution have been set forth under the assumption that competition works faultlessly. Yet attention has been called to the fact that interferences exist, and the effects of such obstructions have been explained. Completely free competition is practically never encountered. There is always some immobility of the factors of production, market information is never complete, and human inertia is always present. Nor could these obstacles be entirely removed even if men strove with the best intentions to destroy them. They are natural — a part of the unescapable conditions of human society. It is sometimes said that in a static state competition would be perfect. If there were no change of any sort men would be able to obtain all the information relevant to their particular tasks, and labor, capital, land, and management would be evenly distributed throughout all industry. With this statement we may agree if we also assume that effects of inertia, impulse, education, and social adaptation are also somehow miraculously abolished. The picture is too nearly perfect. It presents man not as a biologic organism adapting himself to an environment that is both natural and social, but as a machine set to run in grooves. It is a useful device for the first step in economic analysis when we desire, as Alfred Marshall put it, to impound all variables but the one under examination within the pale of other things being equal; but it becomes an obstruction to economic thinking if we refuse to recognize its lack of complete reality.

Natural limitations on free competition prevent the free flow of the factors from place to place and from business to business. Hence these factors do not always earn their potential normal marginal productivity. Some earn more, and some earn less. The position of the entrepreneur enables him to take some of the gain when by good luck it happens that natural obstacles to competition cause the factors which he employs to produce more than the normal rate of return. But he is also compelled to take the lean with the fat; and if the failure of competition works to the disadvantage of the business he controls, he must accept the loss. As we have seen, some of these gains and losses are shared by the laborers, the savers, and the landlords.

Surpluses that arise from these unavoidable limitations on competition, which are not the results of conscious design, are in no sense due to the productive contribution of the entrepreneur. Not only entrepreneurs but also laborers, landlords, and investors participate in surpluses and deficits, caused by these obstacles to complete mobility. Therefore we cannot find in them a peculiar share in distribution. They must be looked upon as deviations of the short-run earnings of the factors from their normal or long-run productivity.

Gains of Monopoly and of Exploitation. Any entrepreneur who can buy raw materials, capital goods, labor, and land in the open market at competitive rates and sell their products in a restricted or partly monopolized market can make gains that are above the normal competitive return of these factors. Legal monopolies of all sorts offer opportunities for this kind of gain. So do partial monopolies due to other causes or combinations that restrict the free flow of capital into the fields in which they operate. But these gains are not due to entrepreneurship. Rather, they result from privileges granted by government, from evasions of the law, and sometimes from sheer fraud and intimidation. In the early history of the railroads large shippers were sometimes able to ruin their competitors by forcing the roads to grant them rates below the authorized schedules which were enforced against their competitors. Today we often hear of "unfair competition" and "unfair methods of getting business." A business man who induces the employees of his rivals to reveal confidential information or who imitates closely the brands of his competitors may gain by doing so, but the reward he receives is not a payment for productiveness any more than is the gain of the confidence man or the swindler. In every field of productive enterprise there are opportunities for shrewd, unscrupulous, and ruthless men to prey upon their associates in business, upon the consuming public, and upon laborers. Their gains are partly wages and partly profits; the latter we may classify as exploitative profits.

In a somewhat different category are the gains from impediments placed by entrepreneurs in the way of the spread of improvements and market information. New processes are carefully guarded to prevent competitors from gaining knowledge of them. The names of profitable patrons, mailing lists, price lists, and the location of desirable markets are regarded as information not to be made public. Probably the significance of these business secrets has been overestimated. In many lines the producers have found that it is desirable that market information should be pooled and then disseminated generally to the entire trade. Certainly it is to the advantage of society to have all advances in technique and in business methods made available to all producers. When the entrepreneur can gain by preventing the spread of improvements or of market information, he may earn more on the factors of production than the going rates placed on them in the open market. But we must decline to call such activities productive.

Summary and Conclusion. Business profits, as the term is commonly used, include the total receipts of the business unit less payments to creditors, to laborers, and to others outside the proprietorship. But business profits are chiefly composed of interest, land rent, and the wages of the owner-operators of the business. The surpluses over and above the normal returns to the factors other than entrepreneurship are pure profits. Pure profits can be analyzed into several unlike constituents: (1) The earnings of unusual ability to manage and to forecast future business conditions, which are returns not clearly distinguishable from wages. (2) Gains due to the uncertainties of economic life, most of which are received by the owners of property. Because the entrepreneur is a property-holder and because of his position in the organization of business, he is in a better position to secure these gains than are laborers or investors who insist on mortgage or other special security. But in the case of the corporation the responsibility for final decisions and risk-taking tend to become separated. When they do, the gains from uncertainty go to the stockholders and not to those who control, except to the extent that the latter are also stockholders. (3) Gains from natural or from legally produced artificial obstacles to free competition. (4) Gains from fraud and other illegal and unethical interferences with economic liberty.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Does the farmer perform entrepreneurial functions? the physician? the owner of preferred stock in a corporation? a hired accountant? the superintendent of a plant who owns no stock in the corporation?

2. Is any part of gain from speculation in stocks economic profit? any part of increase in the value of inventories? If not, how should you classify them?

3. Are exploitative profits earned or unearned incomes? Would production be reduced if the government should take them away? Would the prices of products of the industries in which these profits are found be increased if the government took them?

4. The earnings of especially capable entrepreneurs are sometimes called the "rent of natural ability." In which respects are such earnings like the rent of land? In which respects are they unlike?

5. The entrepreneur has been called the "residual claimant" in distribution. He is said to take the net income that is left after the other factors have been paid their normal shares. Is this statement correct as regards (1) exploitative profits? (2) fortuitous gains? (3) the income from rare natural ability?

6. If you found that a corporation had been earning, on the average, 6 per cent on the original investment of the stockholders, should you say that they had made any pure profit? Suppose that they had earned 20 per cent in one year? 20 per cent on the average?

PART II. INSTITUTIONAL ASPECTS OF DISTRIBUTION

CHAPTER XXIX

TRADE UNIONISM AND LABOR LEGISLATION

The Origin of Trade Unions. Labor organizations are the product of the modern capitalistic economy. They developed first in England, and the following generalizations apply particularly to that country. In the pre-capitalist period families were to a large extent self-sufficient, producing for themselves the things they needed. In so far as they produced for the market, it was a local and restricted market. A The peasant cultivators worked on strips of land which they held under the feudal lord, and they themselves owned the necessary implements and work animals. The handicraftsmen in the towns employed themselves in their own shops and sold their products in the local market. Later came the various inclosure movements, which expelled the cultivators from the soil and made them wageworkers. At the same time the guild system, with its local markets and independent masters, was broken up under the impetus of the growing trade and commerce which sprang from the development of ocean navigation, the discovery of the New World, and the planting of colonies. New industrial towns grew up in which the dispossessed peasants and the former independent masters, who were unable to cope with the new situation, found employment. B But it was not a factory system that they entered: it was a domestic or contract system. The rising group of merchant capitalists who had contact with the distant markets let the work out to be done in the homes of the workpeople. The wage-earners still worked with their own tools, but they worked on materials owned by the merchant capitalists, who sold them in the distant markets. Sometimes the merchant capitalists employed a limited number of workmen in a "manufactory," where they introduced a considerable amount of division of labor, though the work was done with the aid of simple hand tools. The bulk of the work, however, was done under the domestic, or housework, system. The former independent handicraftsmen were unable to compete in the open market with these merchant capitalists. As a result they were gradually forced to give up their own little businesses and sell their labor to the merchant capitalists. In this manner the former peasants and independent artisans gradually became a body of permanent wage-earners. Finally came the invention of machinery and the introduction of the factory system. Under this system the wage-earners no longer worked with their own tools in their own shops or homes, but in factories owned by the growing body of manufacturing capitalists. A similar transformation took place in large part in trade, transportation, and other mechanical pursuits.

Thus modern capitalism created an industrial wage-earning class out of a population which had in large part employed themselves. The transformation can perhaps most vividly be illustrated by showing what would take place if American farming should develop from the small-scale, individual-proprietorship type to the large-scale "factory system" type of farming. Suppose that farming could be more economically carried on by corporations owning large tracts of land and employing scores or even hundreds of workmen and expensive machinery. Should such a development take place, our present small-scale independent farmers would become a body of wage-earners. This is exactly the sort of thing that took place with reference to a considerable body of peasants and artisans of the pre-capitalist era. The modern wage-earning class is a product of the Commercial and Industrial Revolutions. Instead of independent masters owning their own hand tools and selling their own products we have wage-earners with nothing to sell but their labor.

In the Middle Ages the handicraftsmen had formed guilds to protect themselves in the sale of their products. Transformed into wage-earners, they now began to form labor organizations to protect themselves in the sale of their labor. And this became all the more necessary in the latter half of the eighteenth century and the early nineteenth century because of the breakdown of governmental regulations and the introduction of the laissez-faire régime. Formerly products were sold in a local market controlled by custom and guild or governmental regulations. In the new régime the wage-earners found themselves disposing of their labor in an open, competitive market, a market without any of the dikes and bulwarks of organized effort or governmental protection, a market subject to the inundation of a flood of competing labor — woman labor, child labor, prison labor, workhouse labor, immigrant labor, and increasing population. The laborer's problem was to protect himself against this inundation, to protect his labor market. It was out of this situation that trade unionism arose.

How Trade Unions affect Wages.[1] Trade unions are essentially marketing organizations engaged in selling labor services. We have already shown that wages are determined by the marginal productivity of labor; yet it is well-recognized that trade unions may influence the rate of wages. But this fact is in no way inconsistent with the theory of wages which we have presented. We shall briefly summarize at this point the manner in which trade unions may affect wages.

The influence which trade unions have on wages may be stated under two heads: (1) trade unions may assist the wage-earner in obtaining from his employer the full value of his labor; (2) trade unions may alter the value (marginal productivity) of labor.

---

[1] See Alvin H. Hansen's article "The Economics of Unionism," in the *Journal of Political Economy*, August, 1922.

In the first place, a trade union may engage exclusively in the orderly selling of the labor supply of a certain trade. Such a union will admit all comers into the union without any restrictions whatever. It seeks to sell the labor of its members for what it is economically worth — namely, the money equivalent of its marginal productivity. Such a union is an informational and bargaining organization. It keeps in touch with all available facts as to the supply and demand for labor, the condition of the market, the prices at which goods are sold, costs of production, the profits of employers, the rates of wages paid in different localities, and the rates of wages in different industries. The union is thus able to bargain intelligently with the employer. Many an ignorant immigrant laborer has sold his labor far below the going market rate because he knew nothing about the alternative opportunities. He did not know what he could get elsewhere. Only when the union came in and informed him what his labor was worth did he obtain the full value of the marginal product of his labor. One of the important functions of trade unions is therefore to assist workmen in getting wages commensurate with the value of their labor. To do so it is necessary not only to supply information about alternative opportunities, but also to strengthen the reserve or waiting power of the laborers so that it shall not be necessary to take whatever wage is offered. The funds of the union and its power to assess working members constitute such a reserve and strengthen the withholding power of the wage-earners.

In the second place, trade unions may materially affect the market value or, in more technical economic language, the marginal productivity of labor. The marginal productivity of labor may be affected in one of three ways: (1) by changing the personal efficiency of labor, (2) by changing the supply of labor relative to the other factors of production, and (3) by changing the efficiency of the other factors of production.

Employers as a rule contend that trade unions have a tendency to lower the personal efficiency of labor by imposing restrictions on output. Trade unionists usually insist that they do not attempt to restrict output, and that such restrictions as do take place are common to both union and non-union workers. Into this controversy we shall not enter except to say that in Great Britain, at any rate, it appears that union regulations have had the effect of restricting output. During the war, however, British unions agreed to abandon temporarily their restrictive regulations in the interest of larger national productivity. In this country there are indications that some unions at any rate are coming to realize the importance of maintaining production standards.

The effect of trade unions on wages is more apparent, however, in other directions. Trade unions have always sought to limit the total supply of labor. Thus they have opposed child labor, woman labor, prison labor, and immigrant labor. They have had a large influence over our immigration policy. The Knights of Labor, an early labor organization, was largely responsible for the passage of the anticontract labor law which prohibited an employer from advancing the transportation fare to a prospective immigrant on condition that he would work it out after coming to this country. Labor organizations took the lead in the agitation for the exclusion of Chinese labor. The American Federation of Labor has been in the forefront in the long struggle to secure legislation restricting immigration. More recently the incoming of Mexican labor has been a matter of considerable concern to the trade-union movement. Likewise the trade unions in the North are opposed to the northward migration of negro labor. Moreover, the labor movement in all countries is opposed to measures which aim to increase the birth rate. It is said that in the period of the Industrial Revolution in England there were weaving districts in which allowances, over and above the regular wages, were made to employees who had more than two children under ten years of age. Only recently large-family leagues have been organized in France, and numerous funds have been set up by Employers' Associations — notably in France, Germany, and the Netherlands — which pay an allowance (in addition to wages) based on the size of the family. A workman is paid first a wage proportional to the value of the work done, and, in addition, an allowance for his children. The allowance is paid out of a central fund, to which each employer contributes in proportion to the workmen employed. Thus it is immaterial to the individual employer whether he employs single men or family men. It is characteristic that the labor movements of these countries are concerned that these funds shall be so organized that they shall not tend to stimulate the growth of population.

Now it follows necessarily from the principles of wages, which we have elsewhere explained, that if trade unions are able to restrict the total supply of labor, such action will increase the marginal productivity of labor and hence the general level of wages.

But although trade unions are concerned over the supply of labor as a whole, each union is particularly concerned over the supply of labor in its own special trade or craft. The interests of each labor group are affected, it is true, by the general labor market. A low general wage level is a constant menace to a special high-wage group, since there is always a tendency to "spill over" from the flooded market to the more favorably situated one. Nevertheless it is quite possible for a special group to lift its wages far above the general level. Such high wages are to be explained chiefly on the ground of a relative scarcity of labor due to the special skill and training required; but a trade union may increase this relative scarcity in its own specialized labor market through such devices as the closed shop, high initiation fees, limitation of apprentices, and a minimum wage.

The union closed shop means a shop in which only union people can be employed. If this is accompanied by a closed union (a union which refuses to admit new members or at least limits rigidly the number of members) we have a monopoly control over the supply of labor. Here we have, then, "monopoly wages" analogous to "monopoly price." This does not mean, however, that the wages are higher than the marginal productivity of this labor group. Rather it means that this labor group has artificially limited the supply of labor in this line of work to such an extent that the marginal productivity is exceptionally high.

But the great majority of unions do not limit their membership; on the contrary they employ numerous organizers who are constantly seeking to enroll new members. They can in no sense be charged with being monopolistic in the usual sense of that term. Nevertheless, even though the doors of the union are open to all comers, it may still be that entrance can be gained only after paying an exorbitantly high initiation fee. This policy has frequently been practiced by certain unions, though as a rule the entrance fee is not unreasonably high. Or it may be that an unnecessarily long term of apprenticeship is required. Such a policy would decidedly limit the supply of labor in the field. As a rule, however, apprenticeship regulations are not of great importance in modern trade unions.

But all trade unions endeavor, if possible, to establish a minimum-wage scale. If the employers are compelled to obtain all their labor supply through the union, as is the case, for example, in the clothing industry in Chicago, the union may, if it finds it advisable to pursue such a policy, establish a wage scale considerably above the market rate for that grade of work. As a result, however, a portion of those normally following that trade will be unable to obtain employment. The higher the wage the fewer will be the number employed, since the higher-wage scale will necessitate a higher price for the product and as a result the output that can be sold will be smaller. A union may thus find that while it is able to maintain a high wage level, a large number of its members are out of work, and thus the average annual earnings may be no higher than normal. An abnormally high wage scale thus has the effect of limiting the number actually employed in the trade. Hence the wage rate thus arbitrarily established, although in excess of the marginal productivity of the available supply of labor, is not in excess of the marginal productivity of those actually employed.

The extent to which the number will be reduced if the wage scale is pushed above its natural level depends upon the elasticity of the demand for the products of the labor group in question. If clothing-workers receive wages above the natural level, the price of clothing will be higher (assuming constant labor efficiency) and fewer clothes will be sold. Abnormally high building costs reduce the volume of building. In some lines the effect on the volume of employment would be very slight; in others it would be very great.

In the third place, trade unions may affect indirectly the efficiency of the other factors of production and, indeed, often have done so. This is notably true in the case of the management factor. The pressure for high wages exerted by the trade union is a constant goad to managerial efficiency. The employer is forced to buckle down to the job of making industry efficient enough to carry the load of high wages. Mr. E. A. Filene says that if employers are pressed to pay high wages it will help them "to get away from the habit of thinking about lower wages as the only remedy for high cost of production. The big thing in business and industry that makes for high costs is waste — waste growing out of inefficient management and inefficient labor." "In America," says Professor John A. Hobson, "it is generally recognized that the high wages obtainable by all skilled and much unskilled labor in industry operated as the potent stimulus to the great economies in mass production, mass transport, and larger use of machines and mechanical power which distinguish their industrial system."

The Growth of Unionism in the United States. The first nation-wide federation of labor in this country was the National Trades' Union, organized in 1834. This organization sprang up as a result of the upheaval in prices in the thirties and disappeared after the panic of 1837. It was a federation of city centrals, which in turn were composed of local unions in the different trades. The next national federation of labor was the National Labor Union, organized in 1866. During the Civil War upheaval in prices some thirty national trade unions — unions of locals whose members follow the same trade, for example, the National Typographical Union — had been formed along with a large number of city federations. The National Labor Union was a federation of national trade unions and city federations. This organization disappeared in a few years. During the eighties there sprang into prominence an organization known as the Knights of Labor. It became involved in a number of widespread railroad strikes which attracted country-wide attention and brought sharply to the foreground the existence of a labor problem. In the middle eighties a horde of unskilled laborers were inducted into the organization, but after 1886 the membership rapidly declined.

The Knights of Labor, which was a general labor union, soon came into conflict with the American Federation of Labor, an organization built on an entirely different plan. The American Federation of Labor is essentially a federation of national trade unions. Each of these national unions has complete control over its own strike and benefit funds, engages in collective bargaining, and generally regulates its own policies with reference to dues, fees, apprentices, minimum wage, etc. The American Federation of Labor guarantees that 1
each affiliated national union shall respect the jurisdiction of the other affiliated unions so that there will be no overlapping of work or conflict over members. It employs organizers who 2
help the nationals in their organization work, though the main work of this sort is carried on by each national. Moreover, the federation looks after the legislative interests of the labor 3
movement. It seeks to prevent legislation at Washington which is unfavorable to labor and exerts its influence to secure the passage of bills that further the interests of the labor group.

The American Federation of Labor was organized in the eighties, but it did not become conspicuous until the period from 1900 to 1904, when its membership increased sixfold. Another enormous jump in membership came in 1916–1920. A number of strong national unions are not affiliated with the American Federation of Labor. Among these may be mentioned the four railroad Brotherhoods — engineers, conductors, firemen, and trainmen — and the Amalgamated Clothing Workers of America. Nevertheless about four fifths of the total union membership is affiliated with the American Federation of Labor.

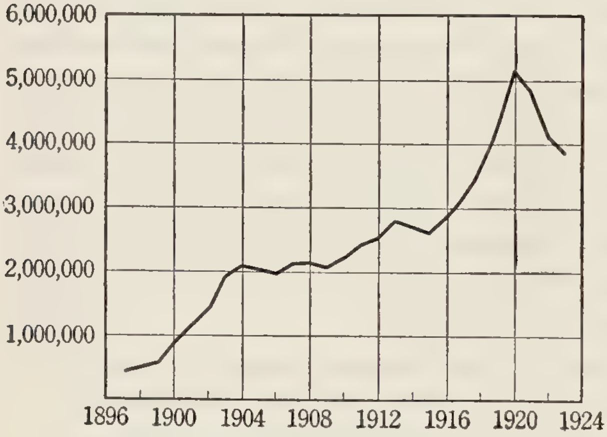

FIG. 38. Union Membership[1]

It has frequently been asserted that less than 10 per cent of the wage-earners in the United States are organized. This statement is based on a misconception of the census figures on occupations. Yet it is not difficult to see how this conclusion was arrived at. There are nearly 4,000,000 organized wage-earners in this country, and there are some 42,000,000 persons gainfully employed. Therefore somewhat less than 10 per cent of the gainfully employed are organized into trade unions. But the term "gainfully employed" includes business men, professional and salaried classes, servants, farmers, and farm laborers. If we subtract these groups we find that there are about 19,000,000 industrial wage-earners, including women and minors. Therefore about 20 per cent of what may properly be called industrial wage-earners are organized at the present time. This represents approximately the strength reached by unionism in England a generation ago. But this figure does not measure the real strength of unionism in America. Unionism is concentrated to a considerable extent in certain strongholds, some of which are of unusual importance. There are five industries in which unionism is particularly strong: transportation, mining, printing and publishing, clothing, and building. In the great fields of manufacture, trade, and agriculture unionism is relatively unimportant. But the five union strongholds constitute strategic points in industry. They are vital to the public in a peculiar way: they are industries that furnish the necessities of life — fuel, clothing, houses, transportation, books, and newspapers. Of all organized labor 25 per cent is in the field of transportation, 23 per cent in building, 11 per cent in mining, 9 per cent in clothing, and 4 per cent in printing and publishing. Nearly three fourths of all organized labor is to be found in these five fields. Yet *all* the wage-earners in these highly unionized fields include not much more than one third of the total industrial wage-earners of the country.

---

[1] Leo Wolman, *Growth of American Trade Unions.*

In all these five fields unionism has secured a powerful foothold. But the struggle for existence is not over. A large percentage of the bituminous coal mines are nonunion. The operating railroad men hold firmly the strong position long since attained, but the railroad-shop crafts which rose to great prominence during the period of government operation have suffered a severe setback. Unionism in the building trades is in ill repute with the public and lost ground in the recent "open shop" drive.

The great steel industry remains almost completely unorganized. In the meat-packing industry, despite the great influence attained by the unions under the bolstering support of the government during the war, the membership dropped close to the pre-war level when this prop was withdrawn. Longshoremen, seamen, boiler-makers, blacksmiths, machinists, and textile workers, after enjoying an enormous rise in membership during the war period, have suffered a terrific decline compared with the high point then reached.

Employers' Associations. Confronting organized labor we have organized employers. Some employers' associations recognize and deal with unions; others seek to destroy or at any rate weaken unions. Among employers' associations which deal with unions may be mentioned the Stove Founders National Defense Association, the American Newspaper Publishers' Association, the United Typothetæ (closed-shop branch), the Building Trades Employers' Association of New York City and similar organizations in other cities, the Illinois Coal Operators Association, and the various railroad systems. Many other employers' associations might be mentioned. The National Council for Industrial Defense was formed to oppose labor legislation. The League for Industrial Rights is especially interested in disseminating information regarding court decisions which deal with certain activities of labor unions, such as the boycott and the sympathetic strike. Of a quite different character is the National Industrial Conference Board, an employers' association interested primarily in research looking toward a more satisfactory adjustment of labor and other problems.

Trade-Union Demands. There are those who profess to believe that unions were once necessary but are no longer so. Unions served, it is said, a useful purpose in the old days, when employers exploited their labor and when wages were miserably low and working conditions intolerable. These people, however, overlook the fact that unions did not arise among the lowest-paid laborers nor among those who suffered the worst conditions. The unions formed in this country a hundred years ago were composed almost entirely of skilled artisans. Their wages were nearly twice as high as the wages of the common laborers. Yet it was the highly paid men that organized. The same was true in England.

Nor has unionism abated with the remarkable rise in real wages that we have witnessed during the last century. We estimate that real wages have increased more than 150 per cent during the last hundred years. Both the skilled and the unskilled have participated almost equally in this rise. It is true that the present wage level for the great mass of unskilled wage-earners is an unsatisfactory one, but it is vastly superior to that of fifty or a hundred years ago. But even though wages were again doubled and trebled, unionism would not disappear. Labor, not unlike other groups, always wants more and more; and particularly is this true so long as we have the existing great disparity in the distribution of wealth and income.

Labor wants not only better wages but more leisure. Here, again, the last century has witnessed a remarkable improvement. A hundred years ago the length of the working day ranged from twelve to thirteen hours, and fifteen to seventeen hours were not unknown. Even as late as 1850 a factory-investigating committee in Massachusetts found the working day to be from eleven to thirteen hours. By 1890 the ten-hour day had become general. At the present time 50 per cent of the wage-earners in manufacturing and mechanical pursuits work eight hours or less, 25 per cent from eight hours to nine hours, and only 25 per cent work more than nine hours. These data only give the length of the standard or basic day; they do not include overtime. The working day has thus been reduced to approximately two thirds of its former length while in the same time real wages have more than doubled.

But labor wants not only good wages and leisure; it also wants security, particularly security of the job. Unemployment is the most serious menace confronting the modern wage-earner, and in this regard we have made little or no progress. Geographical specialization, together with the continued widening of the market, has made the problem of adjusting supply to demand increasingly difficult. Moreover, we have scarcely begun to grapple with the problem of business stabilization. The business cycle is one of the most important causes of unemployment. Seasonal unemployment is also a grave evil, but labor can more easily adjust itself to the regularly anticipated unemployment that comes and goes with the seasons. Not so with the unemployment caused by business depression: not only does this come more unexpectedly, but it is likely to last a considerable period. Moreover, the causes of cyclical unemployment are not readily understood by wage-earners. Workmen are in constant fear that the job will not last. The restrictive policies of trade unionism (such as restriction of output) are in large part the product of the instinctive attempts of wage-earners to make work for everybody. They are crude and ineffective attempts to solve the unemployment problem.

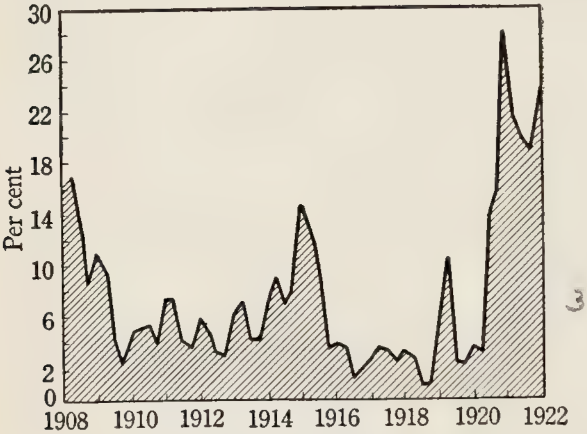

FIG. 39. Unemployment among organized wage-earners in Massachusetts

Modern industry, moreover, gives rise to what may perhaps best be described as transitional unemployment. The continuous introduction 3 of machinery and other labor-saving devices frequently reduces the number of wage-earners employed in the industries affected. The displaced workmen must be absorbed in new industries and new types of service, and this requires effort, ingenuity, and time.[1] Moreover, there is constantly going on a shifting of industries from old regions 4 to new regions and from old products to new products. Shifts 5 in demand and changes in technique account for a considerable volume of unemployment, and the more rapid the change the greater the transitional unemployment that ensues. Thus from 1919 to 1925 manufacturing industries in the United States made a reduction of 600,000 in the number of wage-earners employed, and the railroads reduced their working forces by about 175,000. In the meantime the total volume of production increased considerably in spite of the reduction in the number employed.

---

[1] For a further discussion of this point see pp. 423–427 of this book.

So long as these technical changes take place much unemployment will continue to prevail even though the business cycle were entirely eliminated. In fact, stabilization of the cycle might even have the effect of making it more difficult to absorb those who had become unemployed on account of changes in technique. Were there no longer any boom periods it would be more difficult to start new industries and so absorb the displaced workmen. So long as we have a dynamic society with continual change and progress, a considerable amount of unemployment will necessarily prevail. Unemployment can be minimized, but it cannot entirely be eliminated. Genuine security against unemployment, therefore, makes necessary some form of unemployment insurance. This we shall discuss briefly in a later section.

The Methods of Unionism. The chief methods by which labor organizations exert pressure to achieve their ends are (1) the strike, (2) the boycott, and (3) political influence. The pressure is in proportion to the strategic importance of the strikers as producers, of the boycotters as consumers, and of the political force of a body of voters that will steadily act together. The strike weapon appears to be increasing in importance. The percentage of industrial wage-earners who went on strike each year from 1886 to 1905 was about 60 per cent higher than obtained for the period 1881–1885; from 1915 to 1921 the percentage was more than four times greater than in the period 1881–1885. Since 1922, however, strike movements have greatly declined. Moreover, strikes are becoming more and more national in scope. Railroad strikes and coal strikes affect the industrial life of the entire nation. The boycott, which involves a refusal to buy the goods of an "unfair employer," although not so formidable as it was some years ago, is still an important weapon of labor, particularly in local disputes. Politically, the influence of labor is growing. This influence showed itself notably during the Wilson administration, and it was distinctly felt in the congressional elections of 1922 and in the presidential election of 1924.

Collective Bargaining. The strike and the boycott are the chief weapons used by trade unions to enforce collective bargaining, which is the fundamental goal of unionism. Employers are coming more and more to accept some form of collective bargaining. There are, however, two radically different types: one is the shop-committee, or company, type; the other is the trade-union type. In the former the company deals only with men who are actually in its employ. The employees may choose representatives to deal with the company, but these representatives are themselves employees of the company. No interference by any outside organization is tolerated. Such organization as exists is limited to those actually employed by that company. On the other hand, a trade-union organization is composed of all workers following the same trade in all parts of the country and employed by hundreds or even thousands of different employers. When a trade union engages in collective bargaining with the employer, it does so through the duly elected officers of the union. These men devote all their time to trade-union work and are not in the employ of the company. They are therefore freer and more independent in their negotiations with employers. Moreover, they have back of them a greater reserve strength both in funds and in membership. The whole force of the national organization can thus be brought to bear upon a single employer. The shop committee, on the other hand, has no outside force to rely on.

Employers argue that the shop-committee type is superior from the standpoint of productive efficiency and desirable industrial relations. Since it includes only the employees of a single company, it serves to establish an esprit de corps, a family relationship in the shop so necessary in maintaining efficiency and industrial goodwill. On the other hand, unionists contend that the shop committee leaves labor dependent on and subservient to the employer.

We are touching here upon the essential bone of contention in the "open shop" controversy. The public mind is befuddled more or less with regard to this controversy because of confusion as to the terms used. Strictly speaking, the open shop means a shop in which neither membership nor nonmembership in a union is a necessary condition for employment. It is a shop in which both union and nonunion men may be employed. From this definition it follows that a closed shop is a shop in which one cannot be employed unless he either is or is not a member of a union, depending upon whether the shop is closed to nonunion or to union men respectively. A closed shop, properly speaking, may thus be one in which only union men may be employed, or it may be one in which only nonunion men may be employed; an open shop is one in which either union or nonunion men may be employed. There are industries which run on the open-shop principle in which labor is nevertheless well organized and in which unions deal collectively with the employers. The railroads are a good example. The employer is free to hire a nonunion man, but he must pay him the union wage and give union conditions of work; yet his business is, strictly speaking, an open shop.

Popularly, the terms are used in a very different sense. As the public generally understands the terms, the closed shop means a shop in which the union is recognized and bargains with the employer, whereas the open shop means one in which the union is not recognized by the employer. As popularly used the terms are therefore misleading. Because of this misuse of terms it is difficult for the public to understand that union recognition and collective bargaining are not incompatible with the genuine open shop. Indeed, in order to maintain a real open shop, it is often necessary to secure a union agreement which protects union men in their right to be employed. This confusion of terms has gone so far that employers who insist that their workmen must sign contracts to the effect that they will not join a labor union publicly proclaim that they are running an "open shop."

While there is much talk about the closed shop, the real opposition is for the most part to unionism itself. The great bulk of employers in the United States are still opposed to collective bargaining through trade unions. In Great Britain, on the other hand, the majority of employers take trade-union collective bargaining as a matter of course. But that attitude is a product of a long period of evolution. During the greater part of the first three quarters of the nineteenth century British employers fought the principle of trade-union bargaining as bitterly as American employers do now; but the growing power of labor organizations forced the employers finally to deal with the unions. No doubt many British employers would be glad to free themselves from the limitations imposed upon them by these agreements, but trade unions are too firmly intrenched in the British social order. In the same way, the American employers who now recognize and deal with unions do so not because they prefer this system, but merely because there is no practical alternative.

Unionism in the Modern Industrial Order. The increased growth and influence of labor organizations in all modern industrial countries point to the conclusion that they are inevitable concomitants of the modern industrial order. William Howard Taft has given clear and concise expression to this view in the following sentences: "Organization of labor has become a recognized institution in all the civilized countries of the world. It has come to stay; it is full of usefulness and is necessary to the laborer. . . . There is among employers the bourbon, the man who never learns anything and never forgets anything; the man who says: 'It is my legal right to manage my business as I choose, to pay such wages as I choose, to agree to such terms of employment as I choose, to exclude from my employment union men, because I don't approve of the tenets of the union, and to maintain a family arrangement of my own. I do fairly by my men; I pay them what I think is right, and they will not complain unless some outside union agent interferes. I run a closed nonunion shop and I am happy and propose to continue happy.' This man is far behind in the progress of our social civilization. He lacks breadth of vision extending beyond the confines of his shop. . . . He does not recognize that we have advanced beyond the state in which employers and employees are mere laws unto themselves. . . . But whether we will or not, the group system is here to stay, and every statesman and every man interested in public affairs must recognize that it has to be dealt with as a condition, to be favored in such a way as to minimize its abuses and to increase its utility."[1]

There are, however, many people who would not be willing to accept this conclusion. They believe that the labor problem can be handled better without the trade union. They hold that the trade union brings industrial war rather than industrial peace, and that it is impossible to arrive at any high level of efficiency with labor and management organized in opposing camps. Hence they favor group insurance, welfare work, profit-sharing schemes, sale of company stock to employees, shop committees, and personnel management.

As a matter of fact the question is not so much whether or not such activities as outlined above shall be carried on; it is rather whether these activities shall be dominated and controlled by the employer, or shall be jointly controlled by the employer and the trade union. It might be said, it is true, that joint control could be vested in the employer and a company union functioning through a shop committee. But, as has been pointed out above, the likelihood is that such a committee would be dominated by the employer. Only if they have a strong trade union back of them are the workers likely to retain their independence. Shop committees are, in themselves, not inconsistent with trade unionism and are often (especially in Great Britain) a part of the machinery of trade-union collective bargaining. Only when the shop committee is made a substitute for trade unions, as has so often been true in this country, do the unionists object.

---

[1] Daniel Bloomfield, *Problems of Labor*, pp. 212–214.

There appears to be no way in which wage-earners can make sure of an independent position in their negotiations with employers except through organizations controlled exclusively by themselves. It is for this reason that we believe trade unions are here to stay as far into the future as we can see. There are, indeed, outstanding industrial leaders who have been able "to beat the unions to it" by offering terms of employment superior even to union standards; but the great majority of employers are financially unable to do so even though they would, nor have they the magnetic personality that develops loyalty and enthusiasm for the firm. The enormous turnover of modern industry furnishes adequate proof of this statement. The great bulk of employers drift until the storm is upon them. Industrial goodwill thus has to be worked out through the process of organized groups; through struggle and compromise, and not in the airy realms of a sentimental altruism. Here, then, is the function of the trade union and the employers' association.

Many employers who are foremost in the warfare on unions have themselves had bitter experience with collective bargaining. They know what they are talking about when they point to the restrictive and hampering policies of unions. But these difficulties are not to be wondered at. Industrial government is in its infancy. It will require a long process of trial and error to develop the proper machinery and to produce statesmanlike leaders. Autocracy is simpler and easier than the slower processes of democracy — of consultation and agreement. The evils of unionism are the evils inherent in the democratic method. It is all a part of the gigantic enterprise of making a living under our modern complicated methods of production. Geographical and industrial specialization have made all economic classes highly interdependent. In this complex age consultation and agreement are difficult: they bring grave problems of efficiency and discipline.

The Legal Status of Unionism. From the legal standpoint trade unions are a recognized part of our social order. President Harding in his message to Congress on August 12, 1922, referring "to the warfare on the unions of labor," the "open shop" drive of 1920–1922, said: "The government has no sympathy or approval for this element of discord in the ranks of industry. . . . We recognize these organizations in the law, and we must accredit them with incalculable contribution to labor's uplift. It is of public interest to preserve them and profit by the good that is in them, but we must check the abuses and excesses which conflict with public interest precisely as we have been progressively legislating to prevent capitalistic, corporate, or managerial domination which is contrary to the public welfare."

But while trade unions are lawful organizations in the United States, many of the rights claimed by unionists are seriously questioned by the courts. There have been a number of court decisions which severely restrict the *freedom to organize*. Thus the United States Supreme Court held in Adair *v.* United States[1] that an employer has a constitutional right to discharge a man because he is a union man. This decision thus gives the employer a right to maintain an anti-union closed shop. Again, in Coppage *v.* Kansas[2] the court held that an employer has a constitutional right to force his employees as a condition of employment to sign contracts that they will not join a labor union while they are in his employ. Moreover, in the Hitchman Coal and Coke Company case[3] the court held that it is illegal to try in any way to organize employees if they have signed a contract that they will not join a labor organization.

It is sometimes said that trade unions are not subject to any control in this country; that they are irresponsible bodies, since they are unincorporated. As a matter of fact, trade-union activities are very decidedly regulated in the United States, and in fact labor chafes under the present degree of restriction. While strikes for higher wages and shorter hours are regularly considered legal, there are in the majority of

[1] 208 U. S. 161. [2] 236 U. S. 1. [3] 245 U. S. 229.

states many forms of strikes that are held to be illegal, such as the strike for a closed union shop, a sympathetic strike entered upon for the purpose of aiding some other group of workers to gain their ends, and strikes against the use of nonunion materials. The unions are likewise restricted in the case of the boycott, which is an organized refusal to purchase or handle the products of a merchant or manufacturer. The secondary boycott — that is, a boycott which involves third parties, such as the refusal to buy from retailers who handle the products of the manufacturer against whom the union has a grievance — is in almost all jurisdictions held to be illegal. In many jurisdictions the primary boycott, which does not involve third parties, is held to be legal. Thus in many cities the labor movement advertises through its press and by banners that certain stores are unfair to organized labor, thus seeking to persuade labor sympathizers to withhold their patronage. Such primary boycotts are quite generally upheld by the courts.

Picketing — that is, efforts on the part of strikers to prevent the employer from securing new employees — is generally held to be legal provided it is carried on peacefully and without coercion. Thus, if at each entrance to the plant the strikers have one man carrying a banner which states that a strike is on and urging workers to stay away, such action is generally sustained by the courts. If, however, hundreds of employees are massed in front of the struck plant, such action can scarcely fail to intimidate prospective employees and may amount to coercion. When coercion, intimidation, and violence are in evidence, picketing is generally held to be illegal.

The employer has two main remedies at law if his business suffers unlawful injury at the hands of organized labor. In the first place, he may institute a suit for damages. The most famous suit is probably the one instituted against the Danbury Hatters' Association for injury suffered by the Loewe Manufacturing Company as the result of a nation-wide boycott. In this case[1] damages amounting to about $300,000 were levied against one hundred and seventy-five members of the hatters' union. The members were held liable because they were members of the union whose officers were conducting the boycott.

The decision in the Coronado Coal Company case[2] goes further. It holds that the funds of the union itself as well as the property of the individual members may be seized in payment of damages resulting from the unlawful acts of the union. In the Coronado case, however, no damages were actually assessed against the United Mine Workers of America, since it was found that the national organization had not authorized the strike in which the unlawful acts in question had occurred. Thus the property of the union as well as the property of the individual members may be taken for damages resulting from unlawful acts authorized by the union.

In the second place, the employer may apply to the court for an injunction, which is an order from the court prohibiting unlawful activities that would result in irreparable damage to property. Injunctions seek to prevent acts which would result in loss to the business in question; damage suits seek to recover damages actually sustained.

It is clear that both the strike and the boycott may result in huge business losses even though no physical property is in any way destroyed; but the mere fact that losses are in prospect or are actually sustained is no legal ground in itself for either the issuance of injunction or the sustaining of damage suits. The losses must follow as a result of *unlawful acts*, or the suits for injunction proceedings or damages would not be upheld in the courts. The unions want a broadening of the law so that all forms of strikes and boycotts shall be considered lawful acts. They do not ask that violence or physical injury to persons and property shall be condoned by the law; for any person who commits violence to person or property would still be held in the criminal courts. The unions have no desire to make crimes lawful. They wish, however, to permit wage-earners collectively to withhold their labor and buying power from any person for any purpose whatever, just as an individual may now lawfully do; in other words, they wish trade unions to be freed from the law of conspiracy. This, however, is not possible in the United States without a constitutional amendment, since the Supreme Court has held in Truax *v.* Corrigan[1] that a law granting this freedom is unconstitutional.

---

[1] Loewe *v.* Lawlor, 208 U. S. 274.

[2] 268 U. S. 295.

Government Intervention in Labor Disputes. As strikes have grown in scope and importance the government has increasingly put forth efforts to settle labor disputes peacefully in order to protect the public. Especially has this been true in regard to disputes on the railroads and in the coal mines, industries in which continuity of service is of vital importance to the public. A compulsory-investigation act applicable to the disputes of the railroads was passed in 1888. It remained on the statute books for ten years, but was made use of only once (in the Pullman strike in Chicago in 1894) and then to no effect. In 1898 the Erdman Act was passed providing for mediation and voluntary arbitration. The act was restricted to men actually engaged in the operation of trains. Upon request from either party to the dispute the chairman of the Interstate Commerce Commission and the Commissioner of Labor were to attempt to adjust the dispute by mediation. If that failed they were to urge both parties to submit the dispute to arbitration. Up to 1906 the act was only once made use of and then to no effect. From this time on, however, railroad rates were more rigidly regulated, and hence it became more difficult for the railroads to grant wage increases. From 1906 to 1913 the act was successfully made use of forty-eight times. In no case did either side repudiate an award. In 1913 the Newlands Act was passed. This act provided for the establishment of a Board of Mediation and Conciliation which could act on its own initiative or by request of either party to the dispute. If the board could not mediate the controversy, it was to urge arbitration. Any question arising with respect to the interpretation of an arbitration award was to be referred back to the arbitrators. During the next four years this act was successfully utilized in over seventy cases. But in 1916 arbitration was refused by the unions, and a general railroad strike was imminent. It was averted by the Adamson Act, which granted the basic eight-hour day on the railroads. Late in 1916 President Wilson asked Congress to pass a compulsory-investigation act with compulsory cessation of strike or lockout during the period of the investigation, but no law was passed. Late in 1917 the government assumed control of the railroads. A Railway Wage Board was established and also Boards of Adjustment, which took up all disputes other than those relating to wages. Finally in 1920, when the roads were returned to private control, a United States Railroad Labor Board was established with power to decide all disputes (including wage disputes) by majority vote. However, no penalties except publicity were provided for failure to comply with the award. Publicity was the only penalty which the law provided. The act was, therefore, in no sense a compulsory-arbitration act. Nevertheless the power of public opinion back of the board's award was an important factor in the failure of the shopmen's strike of 1922. In 1926 the Watson-Parker Act was passed, which abolished the Railroad Labor Board and established in its place a Board of Mediation. This act provides for the settlement of disputes by conferences between the interested parties and by voluntary arbitration.

---

[1] 257 U. S. 312.

In the case of the coal miners the Federal government has in recent years frequently intervened in an effort to prevent strikes. In 1919 the Department of Justice secured an injunction prohibiting the officials of the United Mine Workers from distributing strike benefits and compelling them to rescind the strike order. Later the President induced both parties to submit the dispute to a board appointed by the President. Moreover, the Secretary of Labor has at times acted as mediator in coal strikes as well as in industrial disputes, frequently with considerable success.

During the war the National War Labor Board was established by presidential proclamation. No special legal provisions were made for the enforcement of its awards. The appeal to patriotic sentiments was generally sufficient to obtain compliance with its orders. If this was not sufficient, the war powers of the President, by which he could take over essential industries and impose the penalty of loss of industrial exemption, were resorted to. Only in very few industries was it necessary to resort to such drastic measures. Practically, therefore, the War Labor Board had powers of compulsory arbitration during the war period. The board was discontinued, of course, at the end of the war.

Only one American state has tried the experiment of compulsory arbitration. In 1920 Kansas established an Industrial Court with powers of compulsory arbitration in all industries engaging in the transportation or manufacture of fuel, food, and clothing, or in the construction of houses. Men were free to leave any employment; but strikes, picketing, and lockouts were prohibited. The court did not succeed in preventing large strikes, though it served successfully in the settlement of small disputes. In 1923 the Supreme Court of the United States declared that part of the law unconstitutional which pertained to wage-fixing in such industries as meat-packing.[1] In 1925 the Supreme Court held further that hours of labor cannot be fixed in such industries.[2] The net effect of these decisions appears to be to restrict the operation of the act to those industries long recognized in law as public utilities. In 1925 the Kansas Legislature abolished the Industrial Court and transferred such powers as it still retained to the Public Service Commission.

In Canada a compulsory-investigation act with compulsory cessation of strikes or lockout during the period of investigation was passed in 1907. The award of the investigating board is not compulsory. In so far as investigating boards have been applied for, the act has helped to settle disputes, but in the majority of cases it is violated at the outset by refusing to apply for the establishment of a board before a strike or lockout is declared. Penalties for violation of the act have not been enforced.

---

[1] 262 U. S. 522.

[2] 267 U.S. 552.

In Australia compulsory arbitration obtains in most of the states, as well as in the commonwealth itself. New Zealand also has a system of arbitration. Victoria has minimum-wage boards, which set up standard minimum wages for almost all industries. Strikes for wages above the minimum set by the wage boards are not illegal, yet this state has been freer from strikes than some of the states with compulsory arbitration. New South Wales especially has had serious labor disturbances largely because of the turbulent conditions in the coal mines, which are of considerable importance in this province. Compulsory arbitration has not prevented strikes in any of the provinces, though disputes of an interstate scope have in the main been successfully handled by the commonwealth court. Compulsory arbitration has, however, furnished the machinery for negotiation and conciliation between employers and employees. The tendency is more and more to work through joint conferences of the parties to the dispute. Compulsory conference is made use of if necessary. The system may therefore be properly described as legalized collective bargaining.

In the United States the labor movement is strongly opposed to compulsory arbitration. The American Federation of Labor believes it can best attain its ends through trade-union activities rather than through industrial courts or labor boards. It is interested in legislation chiefly to secure freedom from the restrictions imposed by law upon the activities of trade unions. The American Federation of Labor is, however, interested in labor legislation for children and women and in certain special types of legislation for men.

Minimum-Wage Legislation. About a dozen states have passed minimum-wage laws applying to women and minors.

Flat minimum-wage rates are established by law in Arizona, Arkansas, and Utah. In Massachusetts, Minnesota, Wisconsin, Arizona, Washington, California, Colorado, North Dakota, Texas, and Kansas, mimimum-wage boards establish the rate for each industry affected. These boards, on finding by investigation that wages are insufficient in a certain industry to maintain the specified standard of living, may establish a subordinate wage board composed of representatives of the employers, the employees, and the public. These boards, on investigation, recommend the establishment of a certain minimum wage. A public hearing is then held by the permanent wage board or industrial commission, and a scale of wages is finally ordered to be made effective. In most states this award is compulsory, but in Massachusetts the only penalty is publicity in the newspapers regarding firms which refuse to pay the established wage. By a four-to-four decision the United States Supreme Court upheld the Oregon wage law, but in 1923 the minimum-wage law passed by Congress for the District of Columbia was declared unconstitutional.[1] In 1925 the United States Supreme Court refused to sustain the Arizona law, and subsequently a Federal court in Wisconsin held that the law in that state could not be applied to adult females. The Minnesota supreme court has, however, held that the law of that state is valid in its application to minors, and in no jurisdiction has any court held to the contrary. The net effect is that these laws can now be legally enforced only with respect to children and minors.

Hours of Work. Legislation regulating the hours of work is on firmer legal grounds. Thirty-six of our important industrial states limit the hours of labor to eight per day for children under sixteen years of age. About half the remaining states have a nine-hour day. Some of the Southern states still permit eleven hours of work per day for children. Nearly all states limit the hours of labor for women: many to eight or nine hours a day, but the majority place the limit at sixty hours

[1] 261 U. S. 525.

a week. The Oregon law, providing for a ten-hour day for all employees in every mill, factory, or manufacturing establishment, was upheld. The hours of labor are limited for employees of local, state, or Federal governments. On the railroads the hours of labor are regulated in the interest of the public safety. Because of the hazards and dangers to health in the mining industry, almost all our important mining states limit the hours of labor for miners to eight hours a day. On similar grounds a few states have enacted eight-hour laws applying to adult males employed in laundries, electric plants, plaster and cement mills. Ten-hour laws have been applied in certain cases to sawmills, bakeries, brickyards, and textile mills.

Social Insurance. Turning now to social insurance — insurance for wage-earners against accident, sickness, old age, and unemployment — we find legislation in the United States far behind that of European countries. Only in the case of industrial accident is the workingman protected by law in the United States. Practically all states, as well as the Federal government, now have legislation providing for compensation for injuries sustained in industrial employment. The expense is borne exclusively by the employer, who is generally compelled to insure either in a state fund or in a private company. If the employer refuses to come under the compensation act he cannot set up any of the old common-law defenses in a suit for damages, such as (1) contributory negligence of the worker himself, (2) negligence of a fellow employee, or (3) assumption of risk on the part of an employee when he takes the job. The compensation usually amounts to from half the wages to two thirds, with certain upper and lower limits beyond which the compensation cannot go.

In Great Britain the wage-earner is insured not only against accident but also against sickness, unemployment, and old age. In the case of accident the employer alone bears the expense, as in the United States; in the case of sickness and unemployment the expense is borne jointly by the employer, the wage-earner, and the state; in the case of old age the state bears the entire burden.

Many employers in the United States have instituted special funds for their own employees, insuring them against sickness, old age, and even, in some cases, against unemployment. The ladies' garment industry in Cleveland guarantees its workers forty weeks of employment a year and pays unemployment benefits amounting to one half the weekly wage if the employment falls below that level. Recently the men's clothing industry in Chicago, in conjunction with the Amalgamated Clothing Workers, has instituted an employment-insurance plan, supported jointly by the employers and the union, which provides for the payment of unemployment benefits at the rate of 30 per cent of full-time wages for a period not to exceed five weeks per year.

Economic Principles of Labor Legislation. We have now to consider how labor legislation affects the production and distribution of the national income. We shall consider this under two heads: (1) how labor legislation affects the *personal* efficiency (and therefore the marginal productivity) of the wage-earning class; (2) how labor legislation affects in other ways the *economic* efficiency of the labor group.

Much labor legislation obviously has a direct relation to the personal health and efficiency of the wage-earners and their families. Safety and sanitary measures not only reduce the number of days lost during the year but also increase the general health and efficiency of the worker. Similarly, shorter hours will lessen fatigue and thus increase bodily resistance to disease. Moreover, shorter hours will make the workman more alert and energetic during the working hours. Twelve-hour-day workers cannot measure up in hourly output to eight-hour-day workers, and frequently the shorter workday shows an advantage in daily output. Social insurance helps to prevent physical deterioration of the worker and his family when wages cease because of sickness, accident, or unemployment. Minimum-wage laws increase the physical efficiency of the underpaid classes by enabling the worker to obtain the food, clothing, and other necessities of life requisite to a healthful existence.

In so far as these measures serve to increase the personal efficiency of the wage-earning class, their marginal productivity will be increased. Thus in so far as the productivity is increased, the benefits received from these measures are made economically possible and are paid for out of the increased productivity. In the case of minimum-wage legislation the higher wages may thus come out of increased productivity. Social insurance need not result in any diminution of wages, provided productivity is increased sufficiently to pay for the cost of the insurance. If the employer is made to bear the expense of the insurance (as he is in workmen's compensation), it is clear that the wages paid in the long run cannot exceed the *net* marginal product of labor; that is, the amount remaining after the insurance expenses have been met. But the net marginal product need not be lower than before the insurance measure was put into effect if there has been an increase in productivity. In the case of legislation on hours the daily wage need not be reduced if the laborer produces as much in the shorter workday as he does in the longer day.

In the second place, the marginal productivity of labor may be increased by labor legislation even though there is no increase in personal efficiency. The economic productivity of labor depends, it will be remembered, not only on the personal efficiency of the individual wage-earner but also on the relative supply of the different factors of production as well as the efficiency of the factors auxiliary to labor. The minimum-wage legislation may raise the marginal productivity of labor by exerting a powerful pressure on the employer to improve management and equipment to an extent which will make possible the payment of the higher wages required by law. As has been pointed out before, a high wage level is a powerful stimulant to improvements in management and technique.

It is a challenge to the ingenuity of employers, and those who are not able to meet the test are likely to be squeezed out. There is, of course, a limit to such possible improvements. Nor is it possible for all firms to reach the maximum efficiency attained by superior firms. Hence legal minimum wages if pushed very far above the existing level may easily reach a point too high for industry to sustain. In this event, if the law is really enforced (which would be difficult), a considerable number of the wage-earners to whom the law applied would be thrown out of work, since they would not be employed if they could not earn the wages required by law. The *active* labor supply would thus in effect be reduced, and hence the marginal product of labor would be raised to a point sufficiently high to pay the legal wages. In short, the total labor supply could not be employed at the abnormally high wage level.

Naturally the workers first thrown out would be the most inefficient workers. Thus the marginal product would be raised for two reasons: first, because the active labor supply would be reduced; second, because the elimination of the most inefficient would leave the level of personal efficiency higher than before.

A minimum-wage law may thus, for the time being, work a hardship on the least efficient workers, but in the long run even this class may gain. Among the inefficient workers who would be thrown out of employment there might be many children and minors. A longer period for the education of the average child would thus be made possible, and this would increase the efficiency of the next generation of workers. Moreover, all these measures serve to raise workingmen's concept of what constitutes a proper standard of life. The effect of this psychological influence makes itself felt through parents' forethought for the future of their children and probably thus indirectly affects the birth rate. As we have already seen, a restricted birth rate is essential to the attainment and maintenance of a high wage level. Furthermore, the throwing out of work of the most inefficient adult workers would probably compel society to give more liberal aid to defectives and partial defectives with a view to educating and rehabilitating those who can be helped and caring properly for those who are hopelessly deficient. Moreover, legal minimum-wage laws may properly permit lower wage rates to persons who are partial defectives but nevertheless can do useful work and contribute to their own support.

Panaceas and the Labor Problem. There are those who believe that the labor problem will finally be solved, not through trade unionism or labor legislation but through revolution and the complete overthrow of the capitalist employer. For a hundred years and more, thousands and millions of people have firmly believed that the industrial millennium was just round the corner. This is a child's dream; but adults as well as children love to believe in fairy tales. Change is indeed the law of life, and though we are apparently moving in the direction of a sort of socialized capitalism, yet there is no evidence whatever that we shall not always have industrial classes with conflicting economic interests and consequently industrial strife. Modern industry cannot survive without an abundant supply of capital, and it remains still to be seen whether we shall ever be able to produce this supply in sufficient quantity in any other way than through the stimulus of private property. It will by no means be as easy for the world to get along without the private capitalist as it was to push aside that privileged class, the feudal barons of the old régime. Moreover, those who possess the highly technical knowledge which modern industry requires hold a position far more strategic than that occupied by the great mass of wage-earners. Each group is necessary to the other. Economic classes persist because they are rooted in the very nature of the modern complex industrial order and in the ineradicable differences in natural ability. The solution must therefore come not in the supremacy of one class over another, whether capitalist, technician, or wage-earner, but in the coöperation which rests on the firm basis of organized groups.

We used the word "solution" above. There is no solution of the labor problem. It is a part of the eternal problem of human relations: of give and take between people engaged in a common enterprise. The labor problem is, in short, nothing more nor less than the art of living and making a living in coöperation with our fellow men. This is an age-long problem, that will exist as long as there are people struggling to live in a world where production must precede consumption.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Suppose a bricklayers' union so strong that it would be able to raise wages to $25 per day. Point out factors which would come into play making it unlikely that such high wages could be retained.

2. It is possible that a trade union may raise wages without lowering profits or raising prices to the consumer. Under what conditions would this be possible?

3. A trade union may raise wages, and the employers may pass the increased cost of production on to the consumers by charging higher prices. Under what conditions would this be likely to happen?

4. Under certain circumstances it is quite likely that a union by forcing the payment of higher wages would force marginal or typical employers out of business. What are these circumstances? What would be the effect upon consumers? Might such action result in social gain?

5. High wages of union men may result in lowering the wages of nonunion workmen. Under what conditions might this be the case? Would it necessarily be the case? Why not?

6. It is sometimes held that the rate of wages depends exclusively upon the relative bargaining strength of the two opposing parties. What do you think of this statement? Does the bargaining theory answer the question as to where wages come from and what sets the limits to the amount that can be paid in wages?

7. While a union may succeed in fixing wages at a very high level, it cannot force employers to hire the usual number of laborers at the exorbitantly high wage. What might the employers do to make it unnecessary to hire so many high-priced laborers? Might the buying public indirectly force the employer to hire fewer laborers? Explain.

8. Show how it might easily be possible for a certain class of wage-earners to suffer as a result of increasing their daily output. In view of this why does labor as a class generally gain from increased productivity?

9. The British Unemployment Insurance Fund collects approximately $250,000,000 a year. Most of this is paid out in unemployment benefits. A part is used to repay the loan which it became necessary for the government to advance to the fund when the tremendous post-war unemployment wiped out the accumulated surplus. This indebtedness at one time reached about $85,000,000, but it was reduced by the end of 1925 to about $40,000,000. In the United States this scheme has frequently been referred to as a "dole"; but since the contributions to the fund are made mainly by the wage-earners and the employers with only about a 25 per cent subsidy from the state, this term is clearly a misnomer. Discuss the merits and demerits of unemployment insurance.

10. The following table from W. I. King's *Earnings, Hours and Employment in Property and Depression* shows the decline in employment in the United States following the crisis of 1920. Note the industries that are hardest hit.

MAXIMUM CYCLICAL DECLINE OF EMPLOYMENT IN 1920–1922 (PER CENT)

Agriculture	3.18
Extraction of minerals	29.66
Building and construction	18.92
Finance	7.14
Public and professional service	4.57
Domestic and personal service	4.11
Steam railways	29.68
Commerce and trade	2.78
All factories	29.97
Metals and metal products[1]	50.25

[1] Vehicles, railroad cars, and all products not elsewhere recorded are included here.

CHAPTER XXX

DYNAMIC ASPECTS OF DISTRIBUTION

The Wage-Earning Class; Wages. The accompanying chart (Fig. 40) shows the historical trend of money wages[1] in this country for over a century. Certain outstanding movements will at once be noted. Marked fluctuations appear at and after two great wars — the Civil War and the World War. These wage upheavals correspond to the great price upheavals, as will be seen by comparing the wage charts given above with the price chart in Fig. 25, page 340. It is obvious, therefore, that the general price level is one very important factor affecting the money-wage curve. If, however, we disregard these outstanding fluctuations we find that the general trend is steadily upward. No such tendency appears in the general wholesale-price curve. We conclude, therefore, that there are other very important factors operating upon the wage curve besides the movement of prices or the fluctuations in the value of money. If we reduce the wage curve to a purchasing-power basis, the effect of these other forces will become more clearly apparent.

In Fig. 41 is presented the movement of real, or purchasing-power, wages. The money-wage index before 1890 has been reduced to a purchasing-power basis by dividing it by an index of the wholesale prices of food, clothing, fuel, light, and house furnishings weighted roughly according to the expenditures of workingmen's families. This does not give an absolutely accurate index of real wages, even though the underlying data were perfectly accurate and representative. To get an

---

[1] Daily wages are used from 1820 to 1890; weekly wages, from 1890 to 1924. The broken-line curve for 1890–1925 is for the annual earnings of employed workers in manufacturing as given by P. H. Douglas in the *American Economic Review*, Supplement, March, 1926. For the solid-line curve see Alvin H. Hansen's article "Factors Affecting the Trend of Real Wages," in the *American Economic Review*, March, 1925.

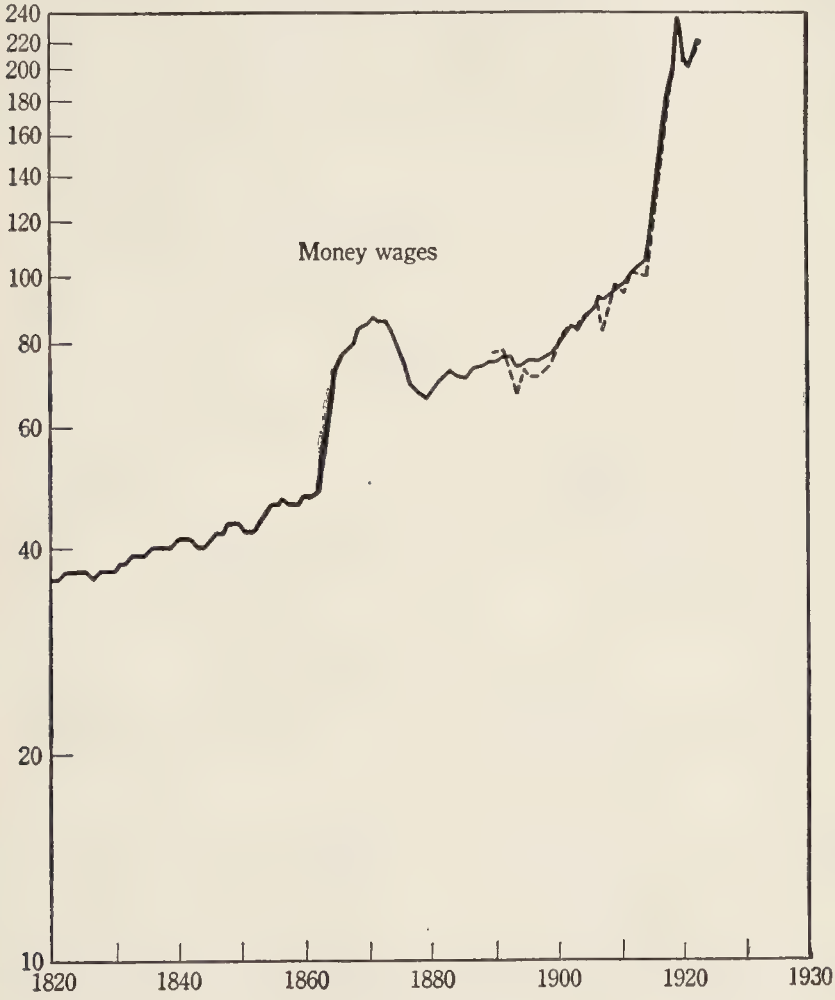

FIG. 40

accurate index of real wages we should, of course, have to reduce money wages to a purchasing-power basis by means of a cost-of-living index. This would include rent as well as food, clothing, etc., and the prices should be at retail and not at wholesale. We have no adequate data on rent before 1913 and we have no continuous retail-price data, except for food, before 1890. Thus before 1890 we have used wholesale prices exclusively; from 1890 to 1913 we have used the retail prices of food and the wholesale prices for other items; after 1913 we have used the cost-of-living index of the Bureau of Labor. According to Paul H. Douglas wholesale prices deviated from retail prices only to a slight degree from 1890 to 1914. However, in periods of pronounced price upheaval wholesale prices fluctuate more violently than retail prices; hence when money wages are reduced to a purchasing-power basis by means of wholesale prices, the real-wage curve is made to fluctuate more violently than the actual facts warrant.

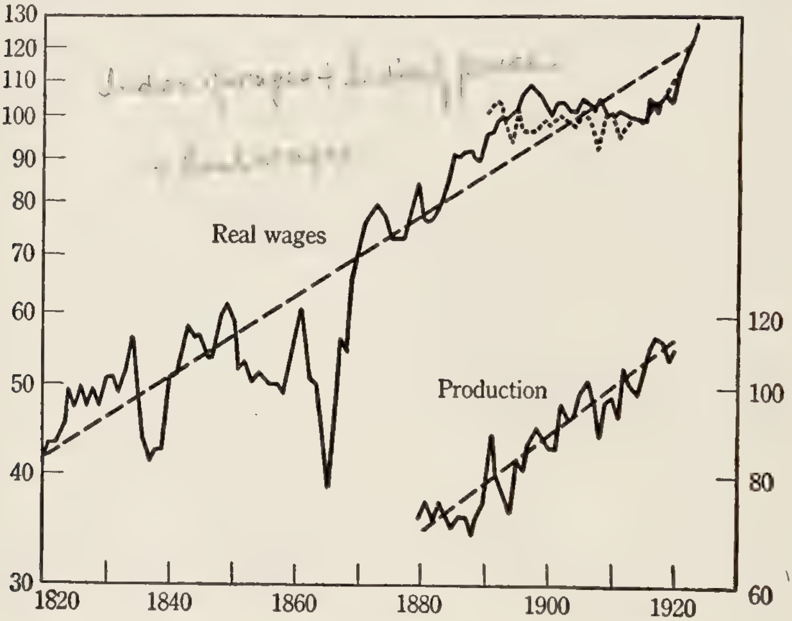

FIG. 41

The Effect of Productivity on Wages. From the chart it is clear that apart from periodic movements, sometimes of longer and sometimes of shorter duration, the long-run trend of real wages has been decidedly upward. This would logically be expected in a century which witnessed a marvelous increase in production. On the chart we show a curve indicating the trend in the production per person gainfully employed.[1] This is by no means the same thing as a curve showing the fluctuations in the marginal productivity of labor. The marginal product of wage-earners is not likely to increase as rapidly as the product per person gainfully employed. As has been fully explained in a previous chapter, wages depend upon the marginal product of labor and not upon the average product.

It is especially important to note the changes in the production of foodstuffs, since food is the most important single item in the workingmen's budget. Real wages are particularly affected by the trend of agricultural production. The per capita production of leading crops showed a marked increase up to the nineties (see Fig. 42). This goes far to explain the rise in real wages in this period. After the nineties the per capita production remained about stationary, with some decline after 1906. The disappearance of free land and the increasing pressure of population, native and immigrant, is one of the factors to be considered in connection with the lag of real wages from 1890 to 1915. Thus the trend in the production of foodstuffs is a factor of especial importance in considering the relation of production to real wages.

But the production of foodstuffs, important as it is, must not be considered alone without regard to general production. Increased productivity in all lines may be expected to increase the real wages of the working class, especially if this increased production reduces the value of goods entering into the work-

---

[1] The leading cause of this increase in productivity per worker is the progressive improvement in the equipment with which each worker is supplied. The primary horse power per worker in manufacturing increased about 90 per cent from 1899 to 1925. The following table gives the increase in the *output per worker* in manufacturing and agriculture from 1899 to 1925 (see *Proceedings of the Academy of Political Science* (July, 1927), p. 7, and the *Conference Board Bulletin* (December, 1927), p. 94):

	MANUFACTURING	AGRICULTURE
1899	100	100
1909	110	106
1919	106	128
1923	134	
1924		147
1925	149	

ingman's budget; but if improvements result in lower cost and increased production of goods not purchased by workingmen, the real incomes of other classes are increased, but there may be no gain for the wage-earners.

Moreover, we have to consider the share of the national product which is distributed through public agencies. Much of the laborer's income does not appear in the pay envelope. An increasing proportion of the national income is being absorbed by the community as a whole, and distributed not according to the principles of value but according to the principles of equality or need. The wage-earner's pay envelope may or may not be affected by these public expenditures, depending on the manner in which the funds are raised. At any rate, these are matters that cannot be overlooked in considering the relation between the movement of wages and national production.

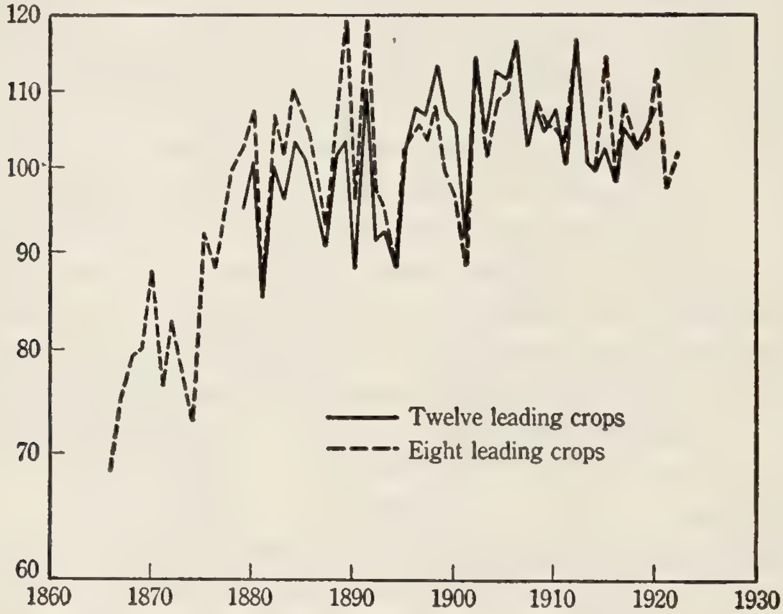

FIG. 42. Agricultural production (per capita)

The Effect of Price Fluctuations on Wages. With the exception of the decade of the seventies, real wages have moved inversely with the general price level. In the burst of prosperity preceding the panic of 1873 real wages rose; in the deep and prolonged depression which followed, labor was so thoroughly deflated that the purchasing power of wages suffered a decline. With the exception of this decade, real wages have risen with falling prices and depression, and fallen with rising prices and prosperity. On the other hand, the purchasing power of labor as a group has risen with prosperity and fallen with depression, because of increased employment in the first instance and widespread unemployment in the second. It is a common fallacy to suppose that the wage-earner's purchasing power will be high if his real-wage rates are high. High real-wage rates do not avail if one is out of a job.

This relationship between real wages and price fluctuations holds true of both the major and the minor movements. Five major movements of prices and wages are apparent. From 1820 to 1849 the tendency of real wages was upward and the trend of prices was downward. From 1849 to 1865 the general trend of prices was upward, and in this period real wages tended downward. From 1865 to 1897 the price trend was heavily downward and the real-wage movement was strongly upward. From 1897 to 1919 the general trend of prices was upward, but real wages remained, on the whole, stationary. With the heavy fall in prices beginning with 1920 real wages rose sharply.

If we center our attention on the cyclical movements we shall notice a similar situation. In the price upheaval of the thirties, which culminated in the panic of 1837, real wages suffered a serious decline. With the rising prices leading to the panic of 1857 real wages again fell heavily. A further drop to still lower levels came with the price upheaval of the sixties. The seventies (the exception to the rule) brought the next period of declining real wages, beginning with the panic of 1873 and culminating in 1877. The last sharp decline came from 1897 to 1900, when wages failed to respond to the unexpected upward turn in prices.

It is evident that price fluctuations have a powerful influence on the distribution of the national income. The marginal product of each factor is no doubt the normal rate of return to each. But this normal distribution is constantly being upset, not only by changes in the relative supply and efficiency of the various factors but also by the redistribution of income accompanying changed prices. If prices rise, the benefit accrues immediately to the entrepreneurial class. In time, however, competition among entrepreneurs compels them to bid up the prices paid for labor, land, and capital until the surplus profits are absorbed. In the meantime a fresh rise in prices has created a new margin of profits above costs. Entrepreneurs naturally do not pass these gains on to the other factors until compelled to do so by the pressure of competition. It is therefore inevitable that, in rising-price periods, wages and other costs should lag behind prices. Rising prices amount in fact to a redistribution of the national income in favor of the entrepreneurial class. They amount to an enforced taxation of wage-earners, salaried persons, investors, and landlords with long-term rent contracts.

From 1897 to 1913, when real wages remained substantially stationary in spite of an enormous increase in national production, business profits far outran the rise in the general price level. This shows itself, for example, in the rise in the stock market. From 1897 to 1913 railroad and industrial stocks advanced nearly 100 per cent. The classical doctrine that profits and wages correlate inversely is abundantly true when applied to the effect of price fluctuations on these distributive shares.

Moreover, this was a period in which huge corporate surpluses were being set aside out of profits. This meant a great addition to our capital equipment. Thus a large part of the increased national production was in the field of producers' goods, as E. E. Day's studies show. This capital equipment came immediately out of the savings of the corporations, but ultimately and in reality out of the enforced savings imposed by the rise in the price level upon wage-earners, salaried people, bond and mortgage holders, insurance-policy holders, etc. Rising prices cause a gap between the marginal productivity of the various factors employed by the entrepreneur and the return that each receives. Indeed, in such periods it is literally true that labor (and certain other factors as well) does not receive the full value of its product.

It may well be, indeed, that certain of the groups enumerated above enjoy a larger income in spite of the enforced saving. Thus though wage-earners find their real wage rates reduced in rising-price periods, yet their real annual earnings are larger because of the increased employment. But that does not obviate the fact that they are none the less taxed by the rise in prices, as is shown by the reduction of their real-wage rates. They get less pay for the work done, but they get larger real incomes by doing more work. Other groups are less fortunate. Salaried people, bond and mortgage holders, and insurance-policy holders suffer a net loss because of the decline in the purchasing power of their fixed money incomes.

On the other hand, labor may get more than its marginal product in falling-price periods. Prices may fall, but labor clings tenaciously to its achieved wage level and does not hesitate even to accept unemployment in preference to wage reductions. Therefore in such periods the employer is hard put to it to reduce costs in some other way than by reducing wages. This he may do by eliminating waste, by introducing improved processes and machinery, and by better management.

Moreover, the marginal product of labor is higher in falling price periods because of the reduction in the labor force. Thus the high real wages are purchased, at least in part, at the price of an excessive volume of unemployment.

The low profits of falling-price periods result in a diminution in corporate surpluses and social saving. The production of consumers' goods assumes a relatively greater importance. Thus, in the falling-price period industry operates to a large extent on the capital accumulated in the previous period of rising prices. Falling prices hand back to the wage-earners (and fixed-income classes) the fruits of the property "stolen" from them by the preceding rising-price level. The "unearned" property acquired by entrepreneurs in the rising-price period is scaled down by bankruptcy, reorganization, and reduction in capitalization in the period of falling prices. Labor now receives, as it were, the "wage dividends," which have been accumulating but which have remained unpaid in the rising-price period.

Labor gains, moreover, at the expense of farmers in periods of falling prices. During the last century and a quarter we have had, along with many movements of less importance, four outstanding periods of severe decline in the purchasing power of farm products, beginning with the following dates: 1817, 1838, 1857, and 1919. These were all periods in which the general price level was falling heavily. What is the reason for this relationship? When general prices fall, industrial depression prevails, and the more precipitate the fall the more severe is the depression. When industry is depressed, the market for farm products is poor for two reasons: first, manufacturing consumes a large part of the raw material produced on farms (80 per cent of the raw material used in manufacture comes from farms) and, second, the purchasing power of wage-earners, much of which is applied to farm products directly or indirectly, is low, because of the prevailing unemployment. In spite of this restricted market, farmers continue to produce approximately the normal supply of the raw materials of industry, while the supply of other factors is rigidly controlled. In a period of depression wage-earners do not go out on the program of selling all their labor at whatever price it will fetch: they prefer unemployment to a precipitate decline in wages. Entrepreneurs curtail production in the face of a falling market. This curtailment checks the extent of the decline in the selling price of manufactured goods. The farmers, however, totally unable to control supply, find themselves in the position of residual claimants. Throwing their commodity on the market in undiminished volume while other groups are

N.B.

controlling supply, they face an adverse purchasing-power ratio. It would be interesting to know how low wage rates would have fallen in 1921 if wage-earners had insisted on selling the entire supply of labor at whatever rates it might fetch.

It is sometimes argued that high wage rates in the depression period are favorable to farmers, for it is thought that the high rates will create a good market for farm products. But unfortunately high wage rates do not help an unemployed man to buy commodities. Widespread unemployment causes low purchasing power in the depression period despite high wage rates. Now labor is probably wise, from a long-run point of view, in resisting wage reductions at all costs, but it cannot be claimed that these high wage rates are favorable to farmers. Labor and the raw materials of farms constitute the chief prime costs of manufactures. They share in the value of the product according to the relative supply of each. If labor curtails its supply (chooses to remain unemployed rather than accept lower wages) while farmers continue to produce at nearly full capacity, the share that goes to the farmer will be inordinately low. Thus we find that in each period of heavily falling prices the purchasing power of wages rises and the purchasing power of farm products falls.

The Effect of Immigration on Wages. The failure of real wages to keep pace with productivity in the two decades preceding 1914 can be explained in part by the unprecedented wave of immigration which began at the turn of the century and continued until the outbreak of the war. In six of the ten years from 1905 to 1914 over a million immigrants arrived annually. The labor market was constantly fed with a new supply of applicants for jobs, making the competitive situation favorable to employers and unfavorable to wage-earners. Under these circumstances it is not surprising that real wages lagged behind the increase in productivity.

On the other hand, the restriction of immigration in recent years has been a potent factor working toward an increase in real wages. Fig. 43 shows the great decline in gross and net immigration during the war period and in subsequent years. European immigration was almost completely stopped by the war, and in the post-war period restrictive legislation has kept the volume at relatively low figures. In 1921 an act was passed, chiefly applicable to Europe, which allowed each country an annual quota equal to 3 per cent of the number of its nationals in this country in 1910. This act which limited immigration from Europe to about 358,000 continued in force until 1924, when a new law was passed which reduced the quotas to 2 per cent of the foreign-born in this country in 1890. Accordingly the total European quota was reduced to about 165,000. This measure especially reduced the immigration from southern and eastern Europe, since relatively few natives of these countries were in the United States in 1890. Under the 1924 act the quota for southern and eastern Europe was only about 20,000.

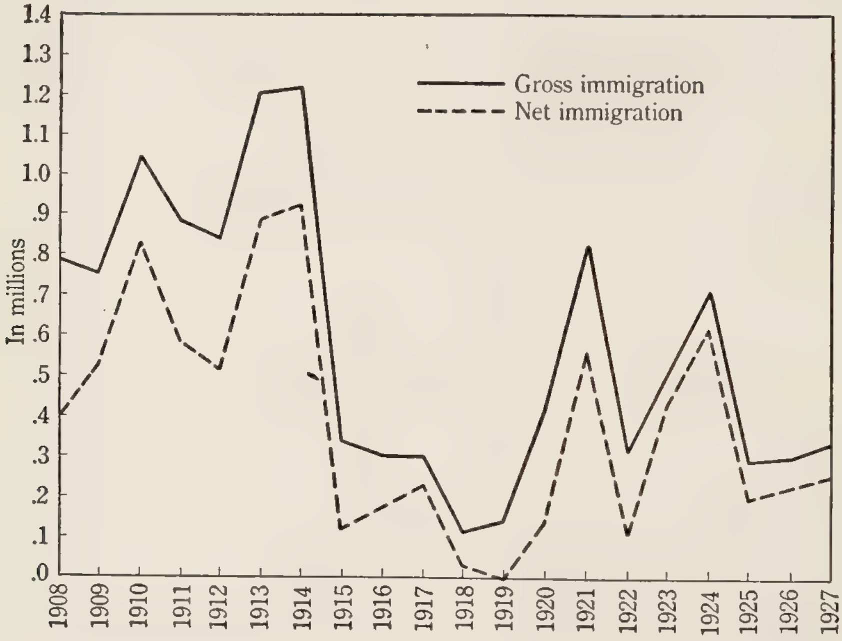

FIG. 43

Table XVII indicates the broad changes in immigration since 1910. The vast net immigration from southern and eastern Europe during pre-war years has been converted to a minus quantity, the outflow to these countries having in recent years exceeded the small number admitted under the quota. On the other hand, the net immigration from northern and western Europe has been reduced less than a third below the pre-war level. The Canadian immigration is three or four times greater than in the pre-war period, and Mexican immigration is two or three times the former level. These large increases are due to the fact that the quota law does not apply to any of the countries of North and South America.

TABLE XVII. AVERAGE ANNUAL NET IMMIGRATION, IN THOUSANDS

PERIOD	MEXICO	CANADA	NORTHERN AND WESTERN EUROPE	SOUTHERN AND EASTERN EUROPE
1910–1914	17	27	157	513
1915–1919	10	57	26	– 3
1920–1924	47	101	110	95
1925–1927	34	89	107	– 13

The Effect of Unionism on Wages. The cost of living has fallen heavily since 1920, but money wages after a moderate decline returned approximately to the peak of the war period. Unions were unquestionably a factor of more or less importance in preventing any drastic reduction in wages during the deflation of 1920–1922. The resistance of the trade unions was particularly evident in certain industries, notably textiles, transportation, mining, clothing, and printing. Moreover, in recent years the skilled and highly organized trades have succeeded in pushing their wage rates to a progressively higher level.

The Farming Class; Land Value and Land Rent. We turn now to a consideration of the fluctuations in the value of farm land. In Fig. 44 we give a curve showing the average price of farm land per acre in the United States (improvements included). The value of land is nothing more nor less than the capitalization of its rental value. However, the capitalization will vary not only with the rental value but also with the rate of interest. Moreover, this curve is by no means an accurate index of changes in land value, since by taking the average price per acre of land in the United States we have a progressive change in the thing evaluated, for new land has continually been added and some old land abandoned. Nevertheless we present the curve for what it is worth; it no doubt presents certain general tendencies.

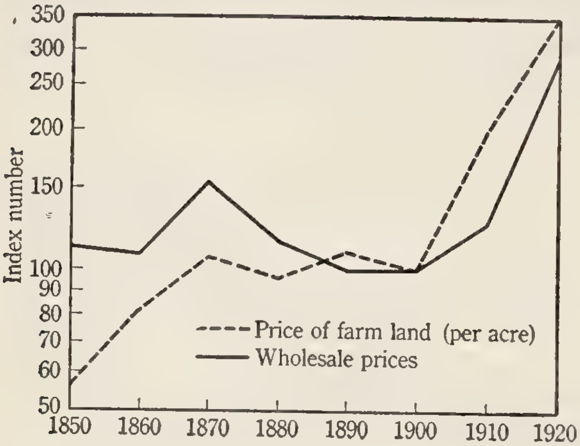

FIG. 44

From 1850 to 1870 the trend of land values was strongly upward, from 1870 to 1900 it was slightly downward, and from 1900 to 1920 there was a phenomenal rise in land values.

The curves (for decade years) in Fig. 44 indicate clearly that a considerable part of the movement of land values is due to the fluctuations in the value of money. Land, like commodities generally, tends to rise as money depreciates in value and to fall with general prices as money appreciates in value.

But this by no means tells the whole story. From 1900 to 1910 land values doubled while general prices increased only about 25 per cent. It is clear that the general increase in prices, or, what amounts to the same thing, the depreciation of money, is not sufficient to explain this phenomenal rise in value. To understand this remarkable increase, we must distinguish between the movement of the prices of farm products and the prices of other commodities. These movements are shown in Fig. 45. It will be noted that the prices of farm products increased 47 per cent from 1900 to 1910, that the prices of other commodities (not including farm products) increased only 18 per cent, and that the general-price index (including farm products) increased only 25 per cent. The disappearance of free land and the increasing pressure of population on limited agricultural resources were obviously having a telling effect. The law of diminishing returns was not being sufficiently counteracted by other factors to prevent agricultural production from falling behind the production of other commodities. Agricultural products were becoming scarce compared with other commodities. Thus the prices of farm products increased more than two and a half times as much in this decade as the prices of other commodities and nearly twice as much as the general wholesale-price index which included agricultural and nonagricultural products.

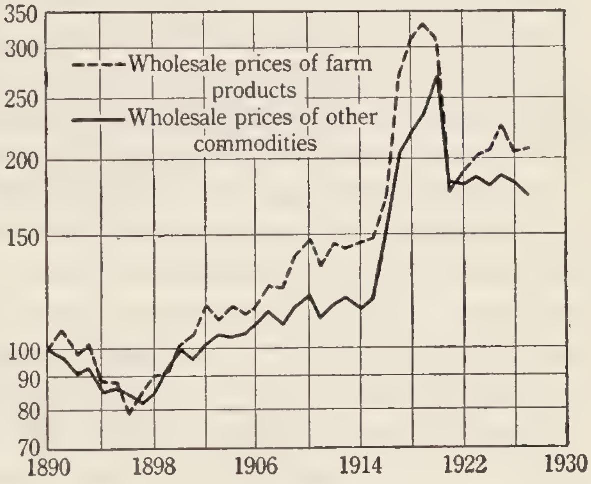

FIG. 45

The effect of these movements on land rents will become clear from an illustration. Let us suppose that a farmer in 1900 sold the products of his farm for $2000 and that his expenses (not including interest on the investment in the land) were $1000. The rental value of his land was then $1000. By 1910 the same crop would sell for $2940. Let us suppose that his expenses increased in proportion to the increase in the general price level.[1] They would then increase from $1000 to

---

[1] The price of farm machinery increased about 15 per cent in this period and the wages of farm labor increased about 40 per cent.

$1250. The net rental value of his land is now $1690. The rental value of his land has thus increased 69 per cent whereas the prices of farm products increased only 47 per cent. It is thus evident that in a period in which the prices of farm products are rising more rapidly than expenses, the rental value of land will increase even more rapidly than the prices of farm products.

But while the rental value of land increased, according to our illustration, 69 per cent from 1900 to 1910,[1] land values doubled. How is this divergence between the rental value and the capitalized value of land to be accounted for? The answer is to be found in the fact that the purchasers of farm land anticipated a continued increase in the rental value of land. The prospective buyer capitalized not only the present rental value of the land but also the anticipated future increase in the rental value. The prospective purchaser of land was therefore willing to pay a price for the land greater than was warranted by the prevailing rental value. And it must be noted that a relatively small annual increase in the rental value accumulates very rapidly. Thus, if the present rental value of a piece of land is $1000 and the anticipated annual increase in the rental value is $50, in twenty years the rental value will be doubled. Now, if we assume a constant rate of interest a $50 annual increase in the rental value if anticipated indefinitely will make the land, now yielding $1000 rent, worth exactly twice as much as if there were no expectation of any increase at all. Anyone who will take the trouble of discounting the anticipated future incomes in the two cases can easily prove this to be the case.

Thus from 1900 to 1910 the capitalized value of land increased more than the rental value because of the anticipated future increases in rental value; the rental value increased more than the prices of farm products because of the fact that expenses did not keep pace with the selling price of the products; and the prices of farm products increased more than the prices of other commodities, because of the law of diminishing returns. Here, then, we find the set of interrelated forces which produced the enormous increase in the value of land from 1900 to 1910. Much the same set of forces was operative from 1910 to 1920. In the latter period, however, land values for the country as a whole did not keep pace with the general rise in prices, though in some states they did. This means that the full increase in rental value was not capitalized, since it was not to be expected that the great rise in the rental value of land during the war period could be counted on to be permanent. Nevertheless people relied on the permanence of this increase to a far greater extent than was justified, as they learned to their sorrow when the drop in farm product prices came in the latter half of 1920. As our curves show, the fall in wholesale prices of farm products was far greater in 1921 than the price decline of other commodities. The prices received by farmers *at the farm* fell still more in proportion. These facts, together with the continuance of high taxes, caused a severe shrinkage of the rental values of farm land, in many cases reducing it to absolutely nothing. There naturally followed a heavy reduction in land values.

---

[1] The increase in cash rents was considerably less than the increase in the economic rent assumed above. Compare C. R. Chambers's *Relation of Land Income to Land Value.*

It is clear from this discussion that one of the important factors affecting the value of land is the movement of the prices of farm products in comparison with the general price level; in other words, the trend of the purchasing power of farm products. The movements of the purchasing power of farm products during the last century will be of interest in this connection.[1]

Table XVIII shows the purchasing of farm products (that is, the prices of farm products in terms of the general prices) by decades from 1790 to 1923. Except for the decades 1820–1829, 1860–1869, and the present decade, the movement has been steadily upward. This rise in the purchasing power of farm products is remarkable and, indeed, somewhat surprising in view of the fact that the nineteenth century witnessed a phenomenal development of agriculture both extensive and intensive. In this century the railroads opened up enormous tracts of virgin land. But manufacturing developed even more rapidly than agriculture. The factory system and large-scale production matched the flood of farm products with a still greater outpouring of manufactured goods. So far as the United States is concerned, agricultural production probably outran the industrial production up to about 1890, but when we consider the world market it is clear that manufacturing production outran agricultural production. After 1890 agricultural production slackened, whereas manufacturing developed apace; consequently the purchasing power of farm products rose sharply.

---

[1] See Alvin H. Hansen's article "The Effect of Price Fluctuations on Agriculture," in the *Journal of Political Economy*, April, 1925.

TABLE XVIII

DECADE	PURCHASING POWER OF FARM PRODUCTS, WEIGHTED INDEX (1913=100)	DECADE	PURCHASING POWER OF FARM PRODUCTS, WEIGHTED INDEX (1913=100)
1790–1799	48	1860–1869	78
1800–1809	51	1870–1879	79
1810–1819	55	1880–1889	84
1820–1829	51	1890–1899	88
1830–1839	62	1900–1909	93
1840–1849	63	1910–1919	104
1850–1859	82	1920–1923	90

For an illustration of these movements, compare the prices of cloth and clothing during the last century with the prices of farm products. As Table XIX shows, clothing was very expensive in the early part of the century and farm products were relatively cheap. As the century progressed, clothing fell in price in relation to farm products, because of the technical progress in the textile and clothing industry.

But the fluctuation in the purchasing power of farm products is not the only factor affecting the farmer's real income. The farmer is a property-owner and, as such, the fluctuations

TABLE XIX

DECADE	WEIGHTED INDEX OF THE PRICE OF FARM PRODUCTS	WEIGHTED INDEX OF THE PRICE OF CLOTH AND CLOTHING
1820–1829	53	171
1830–1839	67	161
1840–1849	59	124
1850–1859	77	122
1860–1869	107	197
1870–1879	87	135
1880–1889	74	104
1890–1899	65	83
1900–1909	81	91
1913	100	100

in the value of money may make or break him. At the present time, for example, it is not the decline in the *purchasing power* of farmers' incomes that is responsible for the foreclosure sales and bankruptcies: it is the high price paid for land at the peak of the recent price movement. Or, to put it in another way, the seat of the trouble is the fact that the farmers' money income (not purchasing-power income) has not maintained its 1919 level, upon which expectancy the land was bought and mortgaged. Even though the purchasing-power income was as great as in 1919, the farmers' difficulty would be nearly as great in many cases as it is today. Overvaluation of land in view of the existing money incomes is the great evil.

From 1865 to 1895 the trend of the purchasing power of farm products was slightly upward, but the money value of farm products was falling greatly. The real income of the farmers was not falling in this period. Yet it was a period of "hard times" for farmers. As property-owners they were being subjected to a constant process of confiscation, accomplished through the constant fall in prices.

The ambition of practically every American gainfully employed in agriculture is to become a landowner, and farmers who are already landowners naturally put their savings into the only kind of investment they know anything about—

namely, land. Rarely is the purchaser of land able to advance the entire purchase price; mortgage indebtedness is therefore the usual lot of farm-owners. Let us suppose that a farmer bought a piece of land for $20,000 in 1867 and that he mortgaged it for $10,000. By 1878 the prices of farm products had fallen to just half the 1867 level. Then there was a considerable rise in the early eighties, followed by a decline to still lower levels in the late eighties and nineties. The trouble was not with the *purchasing power* of the farmer's money income but with the *number of dollars* in his income. The number of dollars had severely declined. His fewer dollars would buy just as much in current goods, but unfortunately the mortgage did not shrink along with other prices. He received fewer dollars for his products and paid fewer dollars for his clothes, horses, machinery, lumber, etc., but he still had to pay the same number of dollars on his mortgage. If the farmer had not been in debt his real income would have been as favorable as in any previous period. Certainly the purchasing power of his products averaged as high as ever before, with the single exception of the prosperous decade of the fifties. But being compelled to pay interest on the mortgage, which refused to shrink along with other money values, his real net income was in fact constantly being reduced, with the decline in the size of his money income. Moreover, this decline in prices brought with it, as has been shown, a decline in the price of land per acre. The farm-owner's equity was being squeezed out. A 50 per cent fall in the price of land would confiscate all the property of a farmer who had paid cash on half the purchase price of his land.

Nor is a period of rising prices altogether an unmixed blessing to the farming population. True, in such periods the mortgages dwindle away relative to the rising money incomes, but here again the evil of overvaluation appears. In such periods land is valued not on the basis of its present income but on the expectation that its rental value will continue to rise. In the period before the World War land was frequently purchased for twice the value warranted by the current rental value. Thus while prospective purchasers — renters who were struggling to climb up into the landowning class — could hope that their money incomes would increase with the rise in prices, the advantage of this was already discounted in the excessive price paid for the land at the outset. It was therefore the rising price level from 1897 to 1920 that caused the overvaluation of land under which farmers now suffer. When the anticipated increase in farm-product prices failed or, worse yet, was converted into a precipitate decline such as 1920–1922 witnessed, the farmer was left stranded high and dry with his highly capitalized land.

Even though the general price level were stabilized, the prices of farm products might still rise because of increasing population pressing on a limited land supply. The expectancy caused by this increase might no doubt result in overcapitalization of land from the standpoint of present earning power, but such an expectancy would at any rate be far more firmly grounded than expectancies based in large part on such arbitrary upheavals in the general price level as we have experienced for one hundred and thirty years. If the general price level were stabilized there would not arise such violent discrepancies between rental value and capitalized value as have harassed farmers for over a century.

The Business Class; Profits. The statistical data with reference to business profits are considerable for the last few years, but do not extend back very far in point of time. With the exception of the banks and the railroads there are no data of any consequence extending over any considerable period of time. We have the index numbers of the prices of stocks back to 1872. Fig. 46 shows the relative prices of industrial common stocks from 1872 to 1927, together with the movement of wholesale prices.[1]

---

[1] By courtesy of Carl Snyder, Federal Reserve Bank, New York City. The figures for 1872–1896 are from the *Commercial and Financial Chronicle*; for 1897–1927, from Dow-Jones and Company.

The wholesale-price trend falls more rapidly from 1872 to 1896 than the stock-market trend. Probably money profits declined more nearly in proportion to the general price level; but interest rates were also falling, and hence the capitalized value of these earnings would decline less rapidly than the money earnings. The upward trend of industrial stock prices from 1896 to 1924 corresponds closely to the upward movement of wholesale prices for the same period. But there are certain divergences that should be noted. It is especially noteworthy that industrial stocks advanced 85 per cent from the low point in 1896 to the period 1905–1913, whereas wholesale prices advanced only 40 per cent. This was clearly a period in which huge profits were being made. Business profits were outrunning the general price level in this period, in part because costs (interest charges, wages, etc.) were lagging behind selling prices. The advantage of these lagging costs accrued to

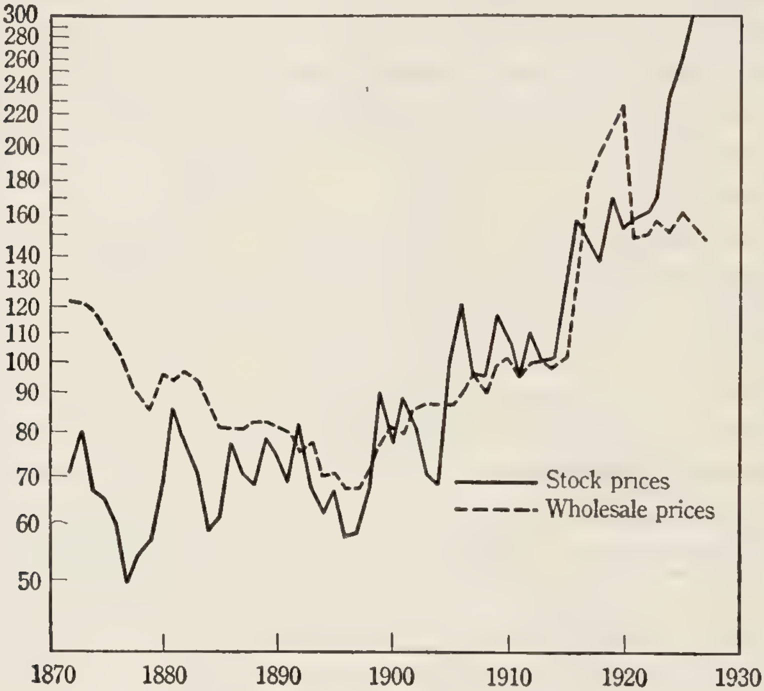

FIG. 46

the entrepreneurial class. Moreover, this was a period of rapid technical progress, and the first effect of such improvements is to enhance pure profits. Furthermore, new industries, such as the electrical industry, the street railway, the telephone, the automobile industry, the motion-picture industry, etc., were developing. New industries, which prove successful, are likely to yield high rates of profit.

At the time of the World War the stock market failed to advance as rapidly as prices. This was partly owing to conservatism in increasing dividends (see Table XX). The total average annual dividends declared during the period from 1916 to 1920 was only 18 per cent in excess of the total dividends declared in 1913. As a result of this low-dividend rate huge surpluses were built up by the larger and more conservative companies. These surpluses helped the better firms to weather the depression of 1921. Profits declined heavily in 1921 (see Table XXI); yet even in this year of pronounced depression the New York Federal Reserve Bank found that less than one third of nearly a hundred large manufacturing and merchandising corporations showed actual losses.

TABLE XX. DIVIDEND AND INTEREST PAYMENTS OF CORPORATIONS IN THE UNITED STATES, IN BILLIONS OF DOLLARS[1]

YEAR	TOTAL DIVIDEND PAYMENTS	TOTAL DIVIDEND AND INTEREST PAYMENTS
1913	0.84	1.78
1914	0.82	1.79
1915	0.79	1.87
1916	0.93	2.13
1917	1.08	2.39
1918	1.02	3.02
1919	0.96	3.19
1920	0.96	3.41
1921	0.92	3.34
1922	0.93	3.40
1923	0.96	3.58
1924	1.01	3.84
1925	1.07	4.08
1926	1.12	4.34

[1] Statistical Abstract of the United States, 1926.

TABLE XXI. NET INCOMES OF CORPORATIONS IN THE UNITED STATES, IN BILLIONS OF DOLLARS[1]

YEAR	TOTAL NET INCOMES OF ALL CORPORATIONS REPORTING NET PROFITS	TOTAL DEFICITS OF ALL CORPORATIONS REPORTING NET LOSSES
1916	8.8	0.7
1917	10.7	0.6
1918	8.4	0.7
1919	9.4	1.0
1920	7.9	2.0
1921	4.3	3.9
1922	7.0	2.2
1923	8.3	2.0
1924	7.6	2.2

Since 1923 large profits have been made despite the high wages paid. The indication is that the productivity per employee has greatly increased. This has been accomplished largely through the introduction of more and better machinery and through better management. Thus from 1923 to 1927 there is no correlation between the movements of stock prices and wholesale prices. During these years there was a phenomenal advance in stock prices, whereas wholesale prices remained practically stationary. This advance in the stock market is to be explained partly by the rapid introduction of cost-reducing improvements, noted above, which resulted in a great increase in business profits, and partly by the large volume of funds available for investment and speculation. These funds came in part from the large profits which were distributed to stockholders and in part from the abundant supply of bank credit which, with our huge gold reserves, our banking system was easily able to provide. The abundance of funds available for investment and speculation progressively lowered the rate of interest.[2] With rising profits and a declining rate of interest, the prices of stocks inevitably rose rapidly. A considerable part of the recent increase in the value of in-

[1] Statistical Abstract of the United States, 1926.

[2] With the rapid increase in incomes we should expect a reduction in the marginal cost of saving.

dustrial shares is thus the result of high profits capitalized at a lower rate of interest. A part is probably due to unwise speculation.

Table XXII shows the pre-war earnings of 10,120 corporations. It will be noted that over 20 per cent earned less than 8 per cent on invested capital (these were marginal firms making no pure profits whatever), that 25 per cent earned from 8 to 12 per cent, and that 30 per cent earned from 12 to 20 per cent. The big bulk of these corporations therefore earned from 8 to 20 per cent. If the 12 to 20 per cent group were divided into two groups, one ranging from 12 to 16 per cent and the other from 16 to 20 per cent, making them strictly comparable to the 8 to 12 group, it would probably appear that the modal group is the 8 to 12 group. A little over 25 per cent earned more than 20 per cent.

TABLE XXII. AVERAGE NET INCOME OF 10,020 CORPORATIONS FOR THE THREE YEARS 1911–1913 [1]

RATIO OF NET INCOME TO INVESTED CAPITAL (PER CENT)	NUMBER OF CORPORATIONS		NET INCOME	
	Actual	Per Cent of Total	Actual (in Thousands)	Per Cent of Total
Under 6	1,368	13.7	$19,011	3.4
6 to 8	737	7.4	26,770	4.8
8 to 12	2,415	24.1	168,246	30.5
12 to 20	2,929	29.3	163,091	29.5
20 to 30	1,366	13.6	93,004	16.8
30 to 40	553	5.5	35,430	6.4
40 to 50	233	2.3	17,817	3.2
50 to 75	259	2.6	16,656	3.0
75 to 100	75	0.7	7,606	1.4
Over 100	85	0.8	4,719	0.9
Total	10,020	100.0	$552,350	99.9

But all corporations do not earn a net profit. An index of business losses may be found in the record of business failures. Table XXIII shows the percentage of business firms failing and the total liabilities of business failures from 1880 to 1927. Table XXIV gives the percentage of corporations reporting

---

[1] David Friday, *Profits, Wages, and Prices*, p. 41.

TABLE XXIII. BUSINESS FAILURES IN THE UNITED STATES, BY DECADES[1]

DECADE	PERCENTAGE OF BUSINESS CONCERNS FAILING ANNUALLY	TOTAL LIABILITIES OF BUSINESS FAILURES ANNUALLY (IN MILLIONS)
1880–1889	0.96	$132
1890–1899	1.10	178
1900–1909	0.90	146
1910–1919	0.88	218
1920–1927	0.95	497

TABLE XXIV. PERCENTAGE OF CORPORATIONS IN THE UNITED STATES REPORTING NET PROFITS AND DEFICITS[2]

YEAR	PERCENTAGE REPORTING NET PROFITS	PERCENTAGE REPORTING NET DEFICITS
1916	60.7	39.3
1917	66.0	34.0
1918	63.6	36.4
1919	65.5	34.5
1920	58.8	41.2
1921	48.0	52.0
1922	55.5	44.5
1923	58.5	41.5
1924	56.6	43.4

net profits and also of those reporting deficits from 1916 to 1924. Of the 236,000 corporations reporting a net income in 1924, 76 per cent earned less than $10,000, 20 per cent between $10,000 and $100,000, and only 4 per cent over $100,000. This latter group, however, earned over 75 per cent of the total net income reported by all these corporations.

The Investing Class; Interest Rates. We next turn to the movement of interest rates. Interest rates on long-time investments are best shown by the yield of high-grade bonds. The bond yields on ten high-grade railroad bonds, reduced to index numbers, are shown in Fig. 47.[3]

---

[1] Statistical Abstract of the United States, 1926.

[2] Ibid.

[3] By courtesy of the Harvard Economic Society; see *Review of Economic Statistics*, 1919 and 1923.

It will be noted that three long-run trends of interest rates from 1890 to 1927 are evident from this chart. From 1890 to 1901 the trend was downward, from 1901 to 1920 it was upward, and from 1920 to 1927 it was downward. To some extent the trends correspond to the movement of the wholesale-price level. Prices, however, reached bottom in 1897, whereas interest rates reached bottom in 1901, four years later. Moreover, the interest rate increased only about 41 per cent from 1901 to 1920 while wholesale prices increased 186 per cent, or over four times as much. Finally, attention should be called to the fact that by 1927 the interest rate had fallen nearly to the pre-war level, whereas prices had been substantially stabilized at 50 per cent above the pre-war level. This illustrates the fact that the rate of interest is not affected by the *level* of prices. While prices were *rising*, interest rates also rose; with the fall in prices in 1921, interest rates also fell. But when prices became stabilized after 1922, interest rates continued to fall until they approached the pre-war normal. This accords with the theory of interest explained in Chapter XXVII. Interest rates are temporarily affected by monetary factors, but in the long run they are controlled by time-preference and by the productivity of capital.

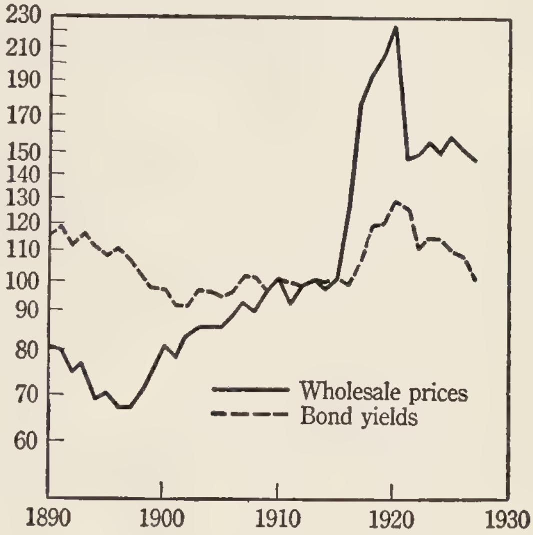

FIG. 47

Our peculiar banking situation since 1914 accounts in part for the great lag of interest rates behind prices from 1914 to 1920 and also for the rapid decline of interest rates from 1920 to 1927. The establishment of the Federal Reserve system increased our banking efficiency, making possible a larger expansion of bank credit on a given reserve of gold. Moreover, during the war period the large gold imports greatly increased the volume of gold reserves, and this tended to retard the advance of interest rates. Similarly, in the post-war period our enormous gold supply has produced a favorable credit situation which has contributed to the recent fall in the rate of interest.

We have noted that while the rate of interest is wholly independent of the *level* of prices, nevertheless with a *rise* in prices interest rates tend also to rise; likewise interest rates tend to fall with a fall in prices. This general relationship is not difficult to explain. A falling-price period is favorable to fixed investors. The stipulated return on the bond or mortgage remains fixed as to the *number* of dollars received at the same time that each dollar received is increasing in value. During such periods a stockholder is likely to receive fewer and fewer dollars, since the money income of business concerns will tend to fall in declining-price periods. Landowners find the money rents from their land rapidly declining. Investments therefore tend to be transferred from stocks and land to the bond market. Hence bonds are bid up in price, and as a result the yield declines.

Conversely, with rising prices, a bondholder is the loser. On the other hand, the earnings on stocks, as we have seen, increase greatly in such periods, as do also the rents from land. Many bondholders readily perceive that it is more advantageous to be a stockholder or a landowner in periods of rising prices. The tendency is to sell the bonds and transfer the investment to stocks or real estate. Bond prices therefore fall in value and, conversely, the bond yields rise.

It is easy, however, to exaggerate the gain or loss enjoyed or suffered by the investor class because of price fluctuations. Let us take an extreme illustration. Suppose a man lends $1000 for a year at 5 per cent. In a year from now he will get back $1050. If no change occurs in the price level in the meantime, the lender will receive back 5 per cent more purchasing power than he lent. This constitutes his interest. Suppose, however, that prices double during the course of the year. It now requires $2000 to equal the purchasing power of $1000 before the rise. To receive back a purchasing power over goods and commodities equivalent to the $1050, on the basis of the old price level, the lender will now have to get $2100. In other words, the interest rate will have to be 110 per cent in order to yield a five per cent return on the real purchasing power lent. However, the investor group as a class does not apply the principal to the purchase of goods and commodities, but rather reinvests it in other loans — bonds and mortgages. Now while commodity prices rise interest rates also rise, and so bond prices fall. Hence in periods of rising prices the purchasing power of the investor's principal is not falling but rising.

Let us suppose that the interest rate doubles along with prices. The interest rate will then rise until at the close of the year, during which prices doubled, the interest rate will be 10 per cent. As a result of this, perpetual bonds (that is, bonds that have no date of maturity) paying a stipulated 5 per cent rate will fall to approximately half their previous price, say from $1000 to $500. Now let us see what will be the situation with our investor who has made a $1000 loan for one year at 10 per cent on the assumption that the price level would rise 100 per cent during the year. He will get back $1100. The $100 is his income from the investment and the $1000 is his principal. Our investor has obviously suffered no loss in real income, since his $100 interest will now (prices having doubled) buy as much as his former $50 annual interest would purchase. Nor has he suffered any loss as a property-holder. He can take his $1000 and buy two one-thousand-dollar bonds (par value), each paying 5 per cent interest. The market value of these two bonds, to be sure, is only about $500 each,[1] because 5 per cent bonds with no maturity date will be worth only about half as much as 10 per cent bonds currently floated.

---

[1] This illustration probably exaggerates the decline of perpetual bonds since it is likely that a future decline in the interest rate will be forecast.

It may appear from this argument that the investor class would neither gain nor lose if the interest rates received by them did indeed rise and fall in proportion to the rise and fall of prices. But, as a matter of fact, such is not the case. The owner of a bond or mortgage receives a fixed money return even though the current interest rate may be changing. If the investor holds a bond or mortgage with a distant maturity, there is no possible way in which he can increase his money income even though interest rates are rising as rapidly as prices. So long as he holds his bond or mortgage he receives, of course, only the former stipulated rate. If, on the other hand, he sells it in order to invest his money in a bond that yields a higher current rate, he finds that he loses on the principal in the same proportion as he gains on the stipulated rate of interest received. As a result, his money income is no higher. If, for example, he is getting 4 per cent and the current rate is 5 per cent, his bond, formerly worth $100 when capitalized at 4 per cent, is now worth only $80 when capitalized at 5 per cent. Eighty dollars invested at 5 per cent will yield no higher return than $100 invested at 4 per cent.

Only in the event that the bond or mortgage matures can the investor take advantage of the higher current-interest rate. If he holds it until maturity he will get back his principal, which may be reinvested at the higher current rate.

We may, then, lay down this general principle: investors, since their income is fixed during the life of the loan, suffer in rising-price periods a loss in real income directly proportional to the fall in the purchasing power of their incomes. In so far, however, as the current rate of interest has risen proportionally with prices, the investor is able to recover his former purchasing-power status every time his loan matures. In point of fact it is, of course, not necessary for him to wait until maturity to recover his purchasing-power status, since the value of a maturing bond depends in part on the date of maturity.

This situation would, however, obtain only so long as prices and interest rates continued to rise in equal proportions.

Should the price level become stabilized at the higher level, interest rates would again fall to the normal level established by the demand for and the supply of capital. But the investor whose bond matures at this juncture finds that he is compelled to reinvest his principal at the lower rate of interest in spite of the fact that prices remain at their former high level. His real income is thus seriously reduced. Moreover, even while prices are in process of rising the rate of interest does not rise *proportionally* with prices. Thus bondholders usually lose heavily in rising-price periods and make large gains in falling-price periods.

Other Factors affecting the Rate of Interest. There are, however, other factors that affect the fluctuations in the rate of interest besides the depreciation of money and the fluctuation in prices. The high rates of contract interest that obtained before 1880 can be accounted for partly by the greater risk that industry faced in a new country with an unstable, fluctuating, and speculative market. Moreover, the country was undeveloped and the demand for additional capital, particularly for buildings and railroads, was very great. The combined influence of the fall in the price level, the stabilization of the market as population and business became more settled, and the decline in the great drafts made upon capital by railroad and building construction, led to the decline in the rate of interest to the low point reached at the close of the nineteenth century. Then the great development of public utilities, the electrical industry, the automobile industry, road-building, modern houses and office buildings, and finally the war, greatly increased the demand for capital, and this, together with the upward movement of prices, led to the great rise in the rate of interest witnessed during the first quarter of the twentieth century.

Probably building construction, including houses, factories, and other buildings, absorbed over one half of our savings for the year 1924; over one tenth went into additional automobiles and trucks; about one fifth went into new railroad equipment and such public utilities as electric light and power, telegraph, and telephone; about one fifteenth went into highway construction and pavements; and about one twentieth into foreign investments.[1]

Altogether we are saving about one sixth of our national income, or approximately from twelve to fifteen billion dollars a year. In spite of this large volume of accumulation the demand for capital appears to be so great that it is not likely that the interest rate will soon fall to the low level reached in the nineties. However, there is always the possibility that the supply of capital will become so great that the rate of interest will fall to a much lower level.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The average weekly earnings of employees in New York State factories in March, 1928, were:

Male shop employees	$34.29
Female shop employees	19.39

2. The following are approximately the wage scales per hour paid in Minneapolis at the present time:

Common laborers	$0.40
Carpenters	1.00
Painters	1.00
Plasterers	1.37
Bricklayers	1.25
Plumbers	1.12
Compositors	0.96
Milk Drivers	0.73
Coal Drivers	0.50

3. Wages paid in manufacturing constituted the following percentages of the value of all manufactured products in the years given:

1909	16.6
1914	16.8
1919	16.8
1921	18.8
1923	18.2
1925	17.1

---

[1] See *The New Republic*, March 11, 1925.

What bearing do these facts have on the degree to which the present high wages encroach on profits? Do they throw any light on the relative efficiency of labor in these periods?

4. The total wage payments made in New York State factories reached the index number of 264 in 1920 compared with 100 for the pre-war period. Average weekly earnings, however, rose only to 222. In 1925 both stood at the figure 223. How can these movements be explained? Why were total wage payments so much higher in 1920 than average weekly earnings?

5. The real wages of clerical and lower-salaried workers are approximately equal to the pre-war level, whereas the real wages of manual workers are from 25 to 30 per cent above the pre-war base. How do you explain these differences?

6. "The secular increase in population would seem to offer every prospect in the long run for an increase in economic rent and in the price of land. If and when this occurs, then the real wages of city workers must be lowered unless such a tendency be counter-balanced by an equivalent or more than equivalent increase in the productivity of manufactures." — DOUGLAS. Argue for or against the proposition that the interests of urban wage-earners and farmers are opposite.

7. "From 1920 to 1924 there was a heavy decline in the membership of trade unions. In this very period there was a rise in real wages. It is therefore evident that trade unions cannot claim any credit for the gain made." Do you agree? Why, or why not?

8. The National Transportation Institute gives the following figures on the relative growth in number of persons employed, in capital investment, and in physical output from 1900 to 1920. The figures are for the year 1920, and are based on 1900 as 100 per cent.

INDUSTRY	PERSONS ENGAGED	CAPITAL	OUTPUT
Agriculture	104	381	138
Manufacturing	208	495	228
Mines	170	302	231
Railroads	199	193	292

Rank the industries according to (1) increase in capital invested per employee; (2) per capita gain in output. Do you see in these facts any application of the principles of distribution?

CHAPTER XXXI

INEQUALITIES OF PROPERTY-OWNERSHIP AND INCOME

We have seen in the preceding chapter that the economic condition of the wage-earning class has greatly improved during the last century. But though the masses are rising in the economic scale, it might well be that the rich are gaining still more. Is property-ownership more highly concentrated in the hands of the very rich now than formerly? Is the middle class being wiped out? Do the poorer classes receive as large a proportion of the total national income now as formerly? How much of the national income goes to the very rich? If these huge incomes were distributed to other classes, how much would they gain? It is not possible to answer these questions with entire satisfaction. We are, however, sufficiently sure of some of the facts to indicate the broad general tendencies.

Inequalities in the ownership of property are in some measure connected with inequalities in income, but not so closely as might at first thought be supposed. As a matter of fact, the distribution of income is much more nearly equal than the distribution of property. There are many professional and salaried persons who earn large incomes but who own little or no property. Moreover, a large part of the national income is paid to wage-earners who own relatively little property. On the other hand farmers, for example, own considerable property but receive relatively small incomes.

The distribution of income is important from the standpoint of consumption, but the distribution of consumption does not necessarily correlate closely with the distribution of income.

In point of fact, many wealthy persons spend only a small percentage of their incomes on personal consumption. Who, then, gets the benefit from these incomes? The wealthy persons in question may spend their surplus incomes on objects for public consumption, such as libraries; they may establish philanthropic foundations; or they may invest in capital goods. Persons with large incomes are therefore in a position to *direct* the currents of public consumption and to *control* production. Inequalities in consumption, of course, follow from unequal distribution of income; but the concentration of control that goes with large incomes and fortunes is probably of even greater significance, for it raises grave questions with respect to democratic institutions and social control. Have these large fortunes and surplus incomes directed consumption and production more wisely than would have been possible with greater democratic control? Does a man like Henry Ford control the consumption of the nation to the extent that he controls its productive processes? Is the democratic control of income and consumption more important than the democratic control of property and production? Is it likely that a democracy can more wisely and successfully socialize its income than its property? Can it achieve genuine control of the one without the other, or are they connected in such a manner that control of the one involves control of the other?

It is not the purpose of this chapter to attempt to give a complete answer to these questions. We propose, rather, to give the student a statistical analysis with respect to the distribution of income and wealth, leaving him, in the main, to deduce such conclusions therefrom as he can.

Distribution of Wealth in Massachusetts. Table XXV shows the distribution of estates among men over twenty-five years of age who died in Massachusetts in the three periods 1829–1831, 1859–1861, and 1889–1891. We have compiled this material from the Twenty-fifth Annual Report of the Massachusetts Bureau of Labor Statistics.

The number of men over twenty-five years of age who die in any one year greatly exceeds the number of estates probated in any one year. The reason for this divergence is that considerably over half the adult males who die have little or no property and hence there is no estate to be probated.

TABLE XXV. DISTRIBUTION OF THE PROBATED ESTATES OF MALES IN MASSACHUSETTS [1]

PERIOD	ESTIMATED [2] PROPORTION OF ADULT MALES WHO DIED LEAVING NO ESTATE TO BE PROBATED (PER CENT)	MINIMUM AND MAXIMUM LIMITS OF CLASS	PERCENTAGE OF TOTAL NUMBER OF ESTATES PROBATED	PERCENTAGE OF TOTAL VALUE OF ESTATES PROBATED	AVERAGE VALUE OF ALL ESTATES PROBATED
1829–1831	52	Under $1,000	47.5	3.5	$4,352
		$1,000 to $5,000	36.6	20.2	
		$5,000 to $25,000	13.6	30.4	
		$25,000 to $100,000	1.9	22.7	
		$100,000 and over	0.4	23.2	
1859–1861	55	Under $1,000	31.7	1.5	8,985
		$1,000 to $5,000	40.9	11.3	
		$5,000 to $25,000	21.6	24.5	
		$25,000 to $100,000	4.3	22.9	
		$100,000 and over	1.5	39.9	
1889–1891	60	Under $1,000	24.6	0.8	13,658
		$1,000 to $5,000	40.9	7.4	
		$5,000 to $25,000	25.9	20.6	
		$25,000 to $100,000	6.2	22.1	
		$100,000 and over	2.3	49.1	

The distribution of wealth among deceased males is, of course, not an absolutely accurate index of the distribution of wealth among the living. It is well known that a considerable amount of wealth escapes the probate courts. Many wealthy men transfer much of their property while they are still alive to members of their family or to charitable, educational, and other similar institutions. All such gifts, transferred during the lifetime of the donor and therefore not included in his will, escape the probate courts. Moreover, life-insurance policies, which are assigned to a designated beneficiary and not to the estate, escape the probate records. Finally a considerable amount of property is held by joint title of husband and wife, and such property passes to the surviving party, without any action in probate court. Attention should also be called to the fact that the age distribution of living adults is, of course, very different from the age distribution of those who die in any given year. The proportion of young men would be much greater among those living. As a rule, men accumulate more property as they grow older. The proportion of men owning property of over $100,000 would probably be less among those living than among those who die in any given year. The average age of deceased males is greater than the average age of those living, and it is likely that the inequality in the ownership of wealth is greater in a given population with the average age of sixty years than in one in which the average age is only thirty years.

---

[1] See the Twenty-fifth Annual Report of the Massachusetts Bureau of Labor Statistics.

[2] Estimated from United States census data on population and mortality.

In 1890 about 60 per cent of the men over twenty-five years of age who died in Massachusetts left no estate to be probated. Of those whose estates were probated, about 25 per cent left less than $1000 worth of property, about 40 per cent left between $1000 and $5000, about 25 per cent left between $5000 and $25,000, and less than 10 per cent left over $25,000.

These figures are likely to give an erroneous impression to the present generation, because the purchasing power of a dollar is today only about 50 per cent of its purchasing power in 1890. A thousand dollars' worth of property today means no more than five hundred dollars' worth of property in 1890. The three periods given in the table are, however, substantially comparable, since the purchasing power of money was not widely different in these periods.

Certain broad conclusions can be drawn from the figures presented in the table. In the first place, it should be noted that in the sixty-year interval from 1830 to 1890 the proportion of men who died leaving no estate to be probated increased from 52 per cent in 1830 to 60 per cent in 1890. Thus there was no indication that the great propertyless masses were becoming property-owners. Rather the tendency appeared to be slightly in the opposite direction.

On the other hand, if we center our attention on that portion of the male population who were owners of property in the three respective periods 1830, 1860, and 1890 we find that the average property-owner in 1860 owned over twice as much as did the average property-owner in 1830 and that the average property-owner in 1890 owned over three times as much as did the average property-owner in 1830; in other words, the property-owners in 1890 were much wealthier than the property-owners one or two generations before.

Now how was the wealth distributed among the different *property-owners*, rich and poor? We are excluding entirely from our consideration here the majority who owned no property whatever except a few personal belongings and who therefore had no estates to be probated; we are centering our attention upon the property-owners, those leaving estates to be probated. What changes took place in the relative property holdings of these men? Do the data show any tendency toward greater concentration of wealth in the hands of a few, or is there a tendency toward greater equality of holdings among all property-owners?

In the first place, it is clear from Table XXV that property-owners were progressively moving up into higher valuation groups, leaving relatively fewer in the lowest group (those owning property worth less than $1000). In 1830 nearly half the property-owners left an estate valued at less than $1000. In 1890 less than one fourth left estates of this value. The rest had moved up into the higher-valuation groups. The $1000–$5000 group increased slightly proportionally, but the upper groups gained the most. The $5000–$25,000 group was proportionally nearly twice as large as in 1830, the $25,000–$100,000 group was over three times as large, and the "over

$100,000" group was over five times as large. This does not mean, however, that property was passing increasingly into the hands of a few rich. What had actually taken place was that property-owners all along the line had become wealthier.

TABLE XXVI. DISTRIBUTION OF TOTAL VALUE OF THE PROBATED ESTATES OF MALES IN MASSACHUSETTS

PROPERTY-HOLDING GROUP, MALES	PERCENTAGE OF TOTAL PROPERTY PROBATED WHICH WAS HELD BY EACH GROUP					
	1829–1831		1859–1861		1889–1891	
Richest group, constituting 1 per cent of estates probated	35		34.5		35	
Moderately rich, constituting 4 per cent of estates probated	22		26.5		26.5	
Upper middle group, constituting 15 per cent of estates probated	25	39	21.0	34.5	23	34
Lower middle group, constituting 30 per cent of estates probated	14		13.5		11	
Poorest group, constituting 50 per cent of estates probated	4		4.5		4.5	

This fact is brought out in Table XXVI. In this table we have more or less arbitrarily divided the property-holders (deceased males with probated estates) into five groups: the richest group constituting 1 per cent of those leaving estates, the rich constituting 4 per cent, the upper middle group 15 per cent, the lower middle group 30 per cent, and the poorest property-owners 50 per cent of the total. It will be noted from the table that the richest 1 per cent owned 35 per cent of the property, both in 1890 and in 1830. There is no evidence, therefore, of increasing concentration in the hands of a few. At the other extreme, the poorest 50 per cent owned substantially the same proportion of the total wealth in 1890 as in 1830. On the other hand, the moderately rich gained relatively at the expense of the middle groups. The rich 4 per cent owned 26.5 per cent in 1890 as against 22 per cent in 1830, whereas the middle groups, constituting 45 per cent of the total, owned 34 per cent in 1890 as against 39 per cent in 1830.

The reader should bear clearly in mind the difference between the *relative* and the *absolute* position of these property-owners. From the absolute standpoint the average value of the property left by members of each of the classifications given was roughly three times greater in 1890 than in 1830; but the gain made by the middle groups was slightly less than that made on the average by all property-owners (the rich gaining slightly more than the average), and the gain made by the richest and poorest groups was almost exactly equal to that made on the average by all property-owners. The general conclusion is that the *distribution* of wealth among property-owners was substantially the same in 1890 as in 1830. On the other hand, the *absolute amount* of property held by each group was about three times greater in the later period.

These conclusions are somewhat modified if we include in our data the estates of both males and females. The data are given in Tables XXVII and XXVIII. It will be noted that the proportion of adults of both sexes who died leaving no estate to be probated *decreased* from 71 per cent in 1830 to 65 per cent in 1890 in Massachusetts, and 57 per cent in 1913–1923 in the United States as represented by twelve widely scattered states and the District of Columbia. It will be remembered, on the other hand, that the proportion of adult males who died leaving no estate *increased* from 52 per cent in 1830 to 60 per cent in 1890. Thus while the proportion of propertyless males was increasing slightly, the proportion of propertyless adults of both sexes was decreasing. Are we to conclude that there were therefore fewer propertyless *families* in 1890 than in 1830? It is difficult to say. It is clear that there was a markedly growing tendency for women to own property. But if this merely means, for the most part, that in 1890 the estate passed to the widow, was held by her for a few years, and was then passed on to the next generation, whereas in 1830 the estate passed directly to the sons, such a situation would not indicate any increase in family well-being. If, on the other hand, an increasing proportion of women, both married and unmarried, held property independently of their husbands or other male relatives, it is clear that this would imply a net gain. Probably both assumptions are true in part, and hence the net conclusion would seem to be that the propertyless class did not change proportionally to any considerable extent either up or down during the period in question.

TABLE XXVII. DISTRIBUTION OF PROBATED ESTATES OF MALES AND FEMALES

PLACE AND PERIOD	ESTIMATED PROPORTION OF ADULTS OF BOTH SEXES WHO DIED LEAVING NO ESTATE TO BE PROBATED (PER CENT)	MINIMUM AND MAXIMUM LIMITS OF CLASS	PERCENTAGE OF TOTAL NUMBER OF ESTATES PROBATED	PERCENTAGE OF TOTAL VALUE OF ESTATES PROBATED	AVERAGE VALUE OF ALL ESTATES PROBATED
Massachusetts, 1829–1831	71	Under $1,000	51.2	4.2	$3,919
		$1,000 to $5,000	34.5	20.9	
		$5,000 to $25,000	12.2	30.6	
		$25,000 to $100,000	1.8	22.8	
		$100,000 and over	0.3	21.6	
Massachusetts, 1859–1861	68	Under $1,000	35.3	2.0	7,694
		$1,000 to $5,000	40.8	25.0	
		$5,000 to $25,000	18.8	23.0	
		$25,000 to $100,000	3.8	37.3	
		$100,000 and over	1.2	12.8	
Massachusetts, 1889–1891	65	Under $1,000	27.1	1.2	10,649
		$1,000 to $5,000	42.4	9.7	
		$5,000 to $25,000	23.7	23.7	
		$25,000 to $100,000	5.1	22.6	
		$100,000 and over	1.7	42.9	
United States, 1912–1923[1]	57	Under $1,000	24.9	0.8	15,428
		$1,000 to $5,000	37.6	6.4	
		$5,000 to $25,000	27.5	19.4	
		$25,000 to $100,000	7.7	22.9	
		$100,000 and over	2.2	50.4	

Other tendencies are substantially the same, whether we take both sexes or males alone. The increase in the average value of estates does not appear so great if both sexes are considered. The upward-climbing tendency of property-owners generally is apparent in substantially the same manner as already discussed with reference to males. Similarly, the proportion of the wealth held by the rich groups, the middle groups, and the poorest group is not appreciably modified by the inclusion of females. The tendencies are the same as have already been indicated.

---

[1] W. I. King, *Journal of the American Statistical Association* (June, 1927), p. 145.

TABLE XXVIII. DISTRIBUTION OF TOTAL VALUE OF THE PROBATED ESTATES OF MALES AND FEMALES IN MASSACHUSETTS AND THE UNITED STATES

PROPERTY-HOLDING GROUPS, MALES AND FEMALES	PERCENTAGE OF TOTAL PROPERTY WHICH WAS HELD BY EACH GROUP			
	Massachusetts (1829–1831)	Massachusetts (1859–1861)	Massachusetts (1889–1891)	United States (1912–1923)
Richest group, constituting 1 per cent of estates probated	34.5	34.5	35	37
Moderately rich, constituting 4 per cent of estates probated . . .	22.5	25.5	25.5	25
Upper middle group, constituting 15 per cent of estates probated . . .	24 ⎫ 39	23 ⎫ 35.5	22.5 ⎫ 35	22 ⎫ 33.5
Lower middle group, constituting 30 per cent of estates probated . . .	15 ⎭	12.5 ⎭	12.5 ⎭	11.5 ⎭
Poorest group, constituting 50 per cent of estates probated	4	4.5	4.5	4.5

Distribution of Wealth in Two Minnesota Counties. Table XXIX gives the property distribution of deceased persons for the year 1923 in Freeborn and Meeker counties, both in Minnesota.[1] These data are fragmentary and cannot, of course, be compared closely with the Massachusetts data. They are, however, interesting illustrations of the inequality in wealth distribution in farming districts in the Middle West. These counties are essentially rural, the largest city being

---

[1] This material was made available by the courtesy of Judge J. W. Wright of Litchfield, Minnesota, and Judge A. W. Johnson of Albert Lea, Minnesota.

Albert Lea, with a population of 11,000. By constructing a Lorenz curve from the data given in the table, we find that the richest 1 per cent owned 8 per cent of the wealth, the moderately rich 4 per cent owned 17 per cent of the wealth, the upper middle group (15 per cent) 34 per cent, the lower middle group (30 per cent) 29 per cent, and the poorest group of property-owners (50 per cent) 12 per cent of the property probated. Moreover, more than half left no property at all so far as the probate records show. From the figures just given it will be noted that the middle groups owned about two thirds of the property probated. This large percentage is to be expected in a distinctly rural community with no large fortunes.

TABLE XXIX. DISTRIBUTION OF THE PROBATED ESTATES OF MALES AND FEMALES IN FREEBORN AND MEEKER COUNTIES, MINNESOTA

PERIOD	PROPORTION OF ADULT MALES AND FEMALES WHO DIED LEAVING NO ESTATE TO BE PROBATED (PER CENT)	MINIMUM AND MAXIMUM LIMITS OF CLASS	PERCENTAGE OF TOTAL NUMBER OF ESTATES PROBATED	PERCENTAGE OF TOTAL VALUE OF ESTATES PROBATED	AVERAGE VALUE OF ALL ESTATES PROBATED
1923	51	Under $1,000	10.0	0.5	$10,809
		$1,000 to $5,000	33.3	7.8	
		$5,000 to $25,000	47.3	52.1	
		$25,000 to $100,000	8.7	33.4	
		Over $100,000	0.7	6.2	

Dr. Willford I. King[1] finds that the distribution of wealth among the living is considerably less unequal than among deceased persons. According to his estimate the richest 1 per cent of the gainfully employed population own about one third of the wealth, and the richest 1 per cent of adult deceased persons (including those with no probated estate) owned over half the wealth of those who have died. The richest 2 per cent of the gainfully employed own about 40 per cent of the wealth, but the richest 2 per cent of adult deceased persons owned 62 per cent of the wealth possessed by those who have died. Ten per cent of the gainfully employed own 64 per cent of the wealth, and 10 per cent of the adult deceased persons owned 86 per cent of the total wealth owned by these individuals.

---

[1] *Journal of American Statistical Association* (June, 1927), pp. 145, 152.

Distribution of Income. Table XXX gives the total income for the people of the United States by decades from 1850 to 1910 and for certain years from 1910 to 1926, together with the per capita income for these years. The real, or purchasing-power, income is also given for the more recent period.

TABLE XXX

	NATIONAL INCOME			PER CAPITA INCOME		
YEAR	KING'S ESTIMATE OF MONEY INCOME (IN BILLIONS)[1]	ESTIMATE OF THE NATIONAL BUREAU OF ECONOMIC RESEARCH		KING'S ESTIMATE OF MONEY INCOME[1]	ESTIMATE OF THE NATIONAL BUREAU OF ECONOMIC RESEARCH	
		Current[2] Money Income (in Billions)	Current[2] Real Income (Index Numbers)		Current[2] Money Income	Current[2] Real Income (Index Numbers)
1850 . .	$2.2			$95		
1860 . .	3.6			116		
1870 . .	6.7			174		
1880 . .	7.4			147		
1890 . .	12.1			192		
1900 . .	18.0			236		
1910 . .	30.5	$28.4	91	332	$307	96
1913 . .		32.0	100		329	100
1918 .		56.0	111		537	103
1920 . .		74.2	114		697	104
1922 . .		65.6	126		597	112
1924 . .		79.4	151		700	129
1926 . .		89.7	165		770	138

[1] W. I. King, *Wealth and Income of the People of the United States.*

[2] "Current income" as used by the National Bureau of Economic Research is "estimated by summating (1) wages, salaries, and pensions, (2) profits withdrawn from business, (3) dividends, interest, and rent received by individuals, (4) the rental value of homes occupied by their owners, (5) interest upon the sums invested in household furnishings, clothing, and the like, and (6) the value of commodities which families produce for their own consumption." "Current income" does not include the savings of business enterprise. See *News-Bulletin of the National Bureau of Economic Research,* February 21, 1927. The figures for the last three years are preliminary estimates.

It is in the division of the income, however, that we are particularly interested. Table XXXI gives the distribution of personal incomes as shown by the Federal income-tax returns for the years 1918 and 1924. It should be noted that the number of incomes in the $1000–$2000 class is much greater than is indicated here, since the income-tax returns for this class include only unmarried persons or married persons not living together. Above the $2000 line all incomes are presumably reported except those subject to special exemption.

TABLE XXXI. INCOME DISTRIBUTION AS SHOWN BY THE FEDERAL INCOME-TAX RETURNS

YEAR	INCOME CLASS	NUMBER OF RETURNS	NET INCOME (IN MILLIONS)	AVERAGE TAX RATE (PER CENT)
1918	$1,000 to $2,000	1,516,938	$2,232	1.19
	$2,000 to $3,000	1,495,878	3,627	0.98
	$3,000 to $5,000	932,336	3,535	2.35
	$5,000 to $10,000	319,356	2,146	4.34
	$10,000 to $25,000	116,569	1,737	8.20
	$25,000 to $50,000	28,542	978	13.32
	$50,000 to $100,000	9,996	680	21.69
	$100,000 to $150,000	2,358	284	33.68
	$150,000 to $300,000	1,514	305	44.64
	$300,000 to $500,000	382	145	54.77
	$500,000 to $1,000,000	178	119	58.65
	$1,000,000 and over	67	137	64.65
1924	$1,000 to $2,000	2,413,881	$3,564	0.29
	$2,000 to $3,000	2,112,993	5,277	0.19
	$3,000 to $5,000	1,800,900	6,828	0.39
	$5,000 to $10,000	437,330	2,991	0.96
	$10,000 to $25,000	191,216	2,855	2.73
	$25,000 to $50,000	47,061	1,600	6.84
	$50,000 to $100,000	15,816	1,067	12.81
	$100,000 to $150,000	3,065	378	20.04
	$150,000 to $300,000	1,876	375	24.69
	$300,000 to $500,000	456	171	26.73
	$500,000 to $1,000,000	242	158	26.87
	$1,000,000 and over	75	156	30.27

Table XXXI includes, however, not much more than 15 per cent of all the income-receivers. To get a more complete picture of the distribution of the income of all the gainfully employed, it becomes necessary to make estimates based on a wide variety of sources. These include interest on tax-exempt bonds, tax-exempt salaries, the imputed rent of home-owners, farm products consumed directly by farmers, annual earnings of the various classes of wage-earners, income of wage-earners derived from sources other than wages, and the incomes of the lower salaried classes. From a wide variety of records the National Bureau of Economic Research has pieced together the available data and constructed an income-distribution table for the year 1918. The accompanying chart (Fig. 48) shows this distribution of income among the various income-receivers in the United States in the year 1918. It includes everyone receiving an income — men, women, and children (with the exception of soldiers and sailors and marines). To some extent it exaggerates the number of persons receiving low incomes, since farmers' incomes are larger than the money

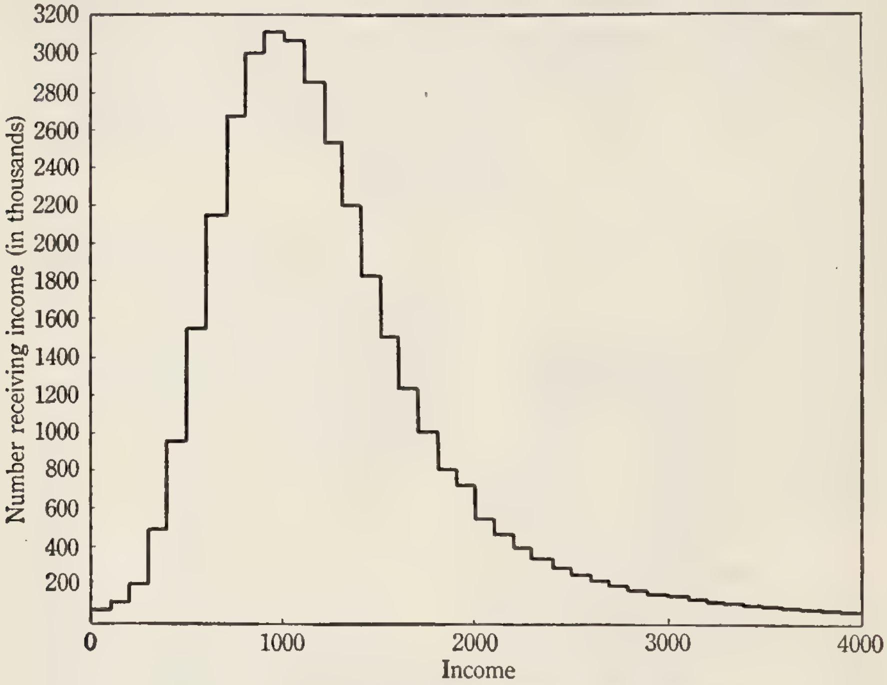

FIG. 48. Distribution of Income in the United States for the year 1918

value here assigned to them. Thus, farm products used by the farmer's family are valued at the price the farmer could get for them rather than at what he would have to pay for them. Moreover, a farm yields a considerable number of perquisites the value of which it is impossible to measure accurately.

The chart gives a picture of the money incomes received by income-receivers, but it does not give an accurate picture of the real distribution of the national income among individuals, for the reason that part of the taxes paid out of certain incomes to the government are redistributed back again by the government to individuals (often not the same as the taxpayers) in various kinds of services. The income of every family is increased, for example, by the services of free schools, free parks, free libraries, museums, paved streets, etc. Certain important taxes, such as the Federal income and inheritance taxes, special assessments, and taxes on property used for business purposes are not deducted from the incomes here given. It follows, therefore, that the net incomes of the well-to-do, who pay these taxes, are smaller than is indicated in the chart and the incomes of the poorer classes are somewhat larger than is indicated, since a part of their income comes to them in the form of free government services.

From the chart it will be noticed that the modal income in 1918 was round $1000. By modal income we mean the income received by the largest number of persons gainfully employed. At first thought it may seem to some that this is contrary to common observation and knowledge. Yet it must be remembered that we are here including all income-receivers, whether men, women, or children. In New York State factories the average weekly earnings in 1918 were $20.35. This would amount to $1058 for fifty-two full weeks. But we all know that few wage-earners work fifty-two full weeks in any year. We cite these figures merely to show the inherent reasonableness of the finding of the National Bureau of Economic Research that the modal income in 1918 was $1000.

It will be seen from the chart that the number of persons receiving the given income rises rapidly up to $1000 and then diminishes until at $4000 the income curve is rapidly approaching the base line. The chart includes only incomes ranging from zero to $4000. If our chart, drawn on the scale here given, were to include all incomes it would have to be extended about a mile and a half to include Henry Ford's income of about $120,000,000 a year. Only about three fourths (78 per cent) of the national income is included in the chart; the remaining fourth (22 per cent) would be included in the narrow but exceedingly long area which we may imagine if the chart were drawn large enough to include the total national income.

Thus 22 per cent of the total national income is included in the imaginary portion of our chart, and this share is received by but 3 per cent of the persons receiving incomes; 28 per cent of the income is included in the area beyond the $3000 line, and this is received by but 6 per cent of the persons receiving incomes; 40 per cent of the total income is included in the area beyond the $2000 line, and this is received by but 14 per cent of the persons receiving incomes. It is clear that the distribution of incomes is very unequal. The height reached by the huge incomes is truly dazzling, viewed from the standpoint of the incomes received by the masses.

Equalitarian Distribution Considered. These facts regarding inequality of income are very impressive. It is no wonder that many people, centering attention upon the enormous incomes of the rich, jump at the conclusion that if only there were more equal distribution of income everyone would be comparatively well-to-do. Let us experiment a little with the possibilities in this direction, having in mind our income chart. Suppose we chop off entirely all the incomes above $4000 and add them pro rata to the incomes below $4000, leaving only an average income for the 3 per cent thus deprived of their incomes. How much would be added by this leveling-down process to the moderate and lower-range incomes? Only 24 per cent. If we chop off all incomes above $3000, leaving only an average income to the 6 per cent thus deprived of their incomes, there would be added to the incomes below $3000 a pro rata increase of only 30 per cent. If we level down the incomes above $2000, leaving an average income for the 14 per cent who would thus be deprived of their income, we would add 43 per cent to the lower incomes on a pro rata basis.

To be sure, an increase of 24, 30, or 43 per cent is not inconsiderable, but it is clear that the dreams of affluence of equalitarians or near-equalitarians would not be realized. As a matter of fact, an addition of 24, 30, or 43 per cent to the real incomes of the lower groups could not possibly be realized; for out of the large incomes above the $3000 and $4000 range come the major proportion of the savings which go annually into new capital equipment as well as a large part of the government revenues. Taxes have in part been deducted from these incomes, since the distribution is estimated in part from the Federal income-tax returns, which permit taxpayers to deduct "personal taxes and all taxes on property not used for business purposes, except special assessments to pay for improvements which benefit property."[1] But other taxes are included, and therefore out of these incomes a considerable part of the taxes have to be paid. In 1922 the sum of $2,226,000,000 was paid in Federal and inheritance taxes alone, and this, of course, was paid out of the higher incomes. Moreover, we are safe in saying that the great bulk of the total savings which go into new capital equipment comes from the higher income levels. Independent estimates for the United States made by King and for the United Kingdom by Stamp indicate that about one sixth of the national income is put aside in savings annually. A considerable part of this comes, it is true, from corporate surpluses, which are not included in the income as here given. Even so, it is evident that a large part of the income of the more well-to-do goes to taxes and capital accumulation. Since these payments would have to be made by somebody, it follows that the addition which could be made to lower incomes by leveling down all incomes above $2000, $3000, or $4000 would be much less than the 43, 30, or 24 per cent suggested above.

---

[1] W. C. Mitchell, *Income in the United States*, p. 139.

It is perfectly clear that there is not much to be gained by the leveling-down process. As a matter of fact, a far greater increase in real income has been achieved by the masses during the last hundred years through increased per capita production than could possibly come through an absolute equalization of incomes, even though we assume that such equalization would have no deleterious effect on national production. It should be noted, however, that if a considerable part of the incomes of the richest class were added to the incomes of the *poorest class alone*, a very material improvement in their status would be effected.

Production versus Distribution. A hundred and thirty years ago two revolutions were sweeping over Europe. The one was a revolution in the methods of production; the other was a revolution in social structure. The first occupied the stage in England; the second, in France. The inventors and production engineers believed that the Industrial Revolution would speedily bring an abundance of material goods within the reach of all mankind at the expenditure of only a few hours of labor a day. The social idealists believed that the French Revolution, with its watchwords of "Liberty, Equality, and Fraternity," would, through its leveling process, uproot poverty. For the first group the chief cause of poverty was insufficient production; for the second group, it was maladjustments in social institutions, inequality of opportunity, and unequal distribution of wealth. Social reform is by no means unimportant, and it is as true today as formerly that eternal vigilance is the price of liberty, but it is none the less true that such gains in material well-being as the last hundred years have witnessed are chiefly the result of inventions and improvements in production. More than any other group it is the inventors and the production engineers who have revolutionized modern life and raised the level of well-being.

Permanency of Relative Distribution of Income. In spite of vast changes in social institutions it appears that in Great Britain and probably in the United States the relative distribution of income has remained substantially the same for a hundred years. Sir Josiah Stamp, in his *Wealth and Taxable Capacity,* gives data on the distribution of all incomes over £200 (about $973 at par of exchange) in 1801 and 1920. Table XXXII gives this distribution.

TABLE XXXII. DISTRIBUTION OF INCOMES IN GREAT BRITAIN IN 1801 AND 1920

PERIOD	INCOME CLASS	PERCENTAGE OF TOTAL NUMBER OF INCOMES OVER £200	PERCENTAGE OF TOTAL INCOMES OVER £200
1801	£200 to £500	61.5	24.0
	£500 to £1000	21.3	18.6
	£1000 to £2000	10.3	17.7
	£2000 to £5000	5.3	20.3
	Over £5000	1.4	19.5
1920	£200 to £500	71.3	29.0
	£500 to £1000	15.8	15.8
	£1000 to $2000	7.8	15.4
	£2000 to £5000	3.7	15.4
	Over £5000	1.3	24.2

Attention should be called first of all to the fact that this table excludes from consideration the great bulk of the population whose incomes fall below £200. Only 5.7 per cent of the income-receivers had incomes above £200 in 1911, and in 1801 the proportion was much smaller; therefore the table includes only the incomes of the relatively well-to-do. But if we take the table as it stands, it appears that a relatively larger proportion of all the incomes over £200 fell into the middle groups (£500 to £2000) in 1801 than in 1920, and that the proportion which fell into the lowest class (£200 to £500) was greater in 1920 than in 1801. Moreover, the richest 1.3 per cent received 24.2 per cent of the income in 1920, whereas the richest 1.4 per cent obtained only 19.5 per cent in 1801. This would indicate that income was somewhat less highly concentrated in 1801 than in 1920. This fact is brought out more clearly in Table XXXIII, which shows the distribution of income as we have estimated it from Stamp's data for the richest 2 per cent, the next richest 8 per cent, the rich 15 per cent, the well-to-do 25 per cent, and the lower 50 per cent of all incomes over £200. From this table it is evident that the richest class received relatively more of the income in 1920 than in 1801. The difference is not great, however.

TABLE XXXIII. DISTRIBUTION OF ALL INCOMES OVER £200

INCOME-RECEIVING CLASS	PERCENTAGE OF TOTAL INCOMES OVER £200 RECEIVED BY EACH CLASS	
	1801	1920
Richest class, constituting 2 per cent of all incomes over £200	22.5	28.5
Next richest class, constituting 8 per cent of all incomes over £200	25.0	22.5
Rich class, constituting 15 per cent of all incomes over £200	18.5	17.0
Well-to-do class, constituting 25 per cent of all incomes over £200	16.0	14.5
Lower half, 50 per cent of all incomes over £200	18.0	17.5

Stamp calls attention to the fact that incomes were not as fully reported in 1801 as in 1920, and it is possible that there was relatively greater evasion in the lowest income group — that between £200 and £500. If we exclude this group we have for incomes over £500 the distribution shown in Table XXXIV.

From Table XXXIV it will be seen that the distribution of incomes above £500 was closely similar in the two periods, but again it is evident that the richest class received a larger percentage of the total income in 1920 than in 1801. This fact can be seen more clearly if we estimate the proportion of the total income received by the richest 5 per cent, the next richest 15 per cent, the rich 30 per cent, and the lower 50 per cent of

TABLE XXXIV

PERIOD	INCOME CLASS	PERCENTAGE OF TOTAL NUMBER OF INCOMES OVER £500	PERCENTAGE OF TOTAL INCOMES OVER £500
1801	£500 to £1000	56.0	24.4
	£1000 to £2000	26.3	23.2
	£2000 to £5000	13.9	26.7
	Over £5000	3.8	25.7
1920	£500 to £1000	55.2	22.3
	£1000 to £2000	27.3	21.8
	£2000 to £5000	13.0	21.8
	Over £5000	4.5	34.1

all incomes over £500. These data are given in Table XXXV. Here again it appears, as in Table XXXIII, that the relative gain of the richest class has come chiefly at the expense of the next richest classes and not at the expense of the lowest-income group.

TABLE XXXV. DISTRIBUTION OF ALL INCOMES OVER £500

INCOME-RECEIVING CLASS	PERCENTAGE OF TOTAL INCOMES OVER £500 RECEIVED BY EACH CLASS	
	1801	1920
Richest class, constituting 5 per cent of all incomes over £500	29.5	35.5
Next richest class, constituting 15 per cent of all incomes over £500	26.0	22.5
Rich class, constituting 30 per cent of all incomes over £500	23.5	20.5
Lower half, 50 per cent of all incomes over £500	21.0	21.5

While the relative distribution of incomes appears to be substantially similar in 1801 and in 1920, the *proportion of the population* included in the income levels given was very much greater in 1920 than in 1801. Stamp reaches the following conclusion: "The total nominal income has increased much more than the total population — the increase has surged upwards through all the fixed classes, so that there is a smaller population in the ranks of the poorest, with a nominal income of say under £80 a year, and many more in the over £5000 class, but the *slope* of the distribution, that is, the relation between one section or class and another, has hardly altered."

There is thus an astonishing similarity in the distribution of income in two periods separated by a century and a quarter, during which time revolutionary changes had occurred in economic and political institutions. Similarly, Professor A. L. Bowley in his *The Change in the Distribution of the National Income, 1880–1913* concludes that the increase in the per capita income (about 35 per cent) in the period in question was shared with remarkable equality among all the various classes. This suggests the possibility that the dominant forces controlling the distribution of income have hitherto been deeper and more fundamental than mere social institutions. It would seem to point to a "fixed system of causation" such as the great inequality of native capacities in the human race. This explanation overlooks the important fact, however, that throughout the period in question the basic institutions of private property and the inheritance of private property have prevailed intact. The permanence of these fundamental institutions, and the relatively fixed distribution of inherited capacities, probably account in large part for the persistence of the scheme of income distribution which has prevailed in modern communities during the last century or more.

That the distribution of income does not conform to the distribution of inherited capacities is clear from an examination of Fig. 48. The income-distribution curve, if enlarged to include all incomes, would be lopsided to an extreme degree; on the other hand, it is likely that a curve showing the distribution of ability would be fairly symmetrical. Above all, it is the institutions of private property and the inheritance of private property which distort the income curve. Indeed, the lower part of the income curve, between zero and $2000, is fairly symmetrical, and these incomes are derived largely from personal services and only to a small extent from property.

Yet it is probable that even though all income were earned from personal services, the income curve would still be skewed to the right, though certainly much less so than under existing property institutions. The artist who is just a little better than his nearest rival can command a much larger price for his paintings. The physician who makes a great reputation may be barely, if at all, superior to his competitors, but he will be able to demand much larger fees for his services. The situation appears to be not altogether unlike a race in which the winner obtains a large prize and his nearest competitor receives nothing at all.

Property Institutions and Social Utility. In so far as inequality in income distribution is due to the inheritance of private property no justification for such inequality can be found on the basis of personal merits and demerits.[1] The person who inherits property, such as natural resources or capital goods, receives an income for the services of his property. The reward is indeed for service rendered, but it is a *property service* and not a *personal service*. Distribution, in the existing order, is, indeed, based on the principle of reward according to the service rendered, but under the prevailing property institutions the *service rendered* may have no relation to *personal deserts*.

How, then, can we justify private property, especially the inheritance of private property, and the inequalities to which they give rise? They can be justified only on the ground of *social utility*. It can scarcely be questioned that inequality of income has in a marked degree facilitated the rapid accumulation of capital which we have witnessed during the last century. The accumulation of capital has contributed greatly to such economic progress as we have made during this period.

---

[1] An important exception to this statement is found in those cases in which the children have worked in their father's business without adequate financial return by way of wages, educational opportunities, or some other form of reward for the services rendered.

Without it we could never have reached the per capita income which we have attained. Inequality of income has thus been one of the factors that has enabled us to raise the *level of income* all round, but it has left us with the same degree of inequality. Other aspects of the social utility of private property will be discussed in the chapter on "Socialism."

But with the wider diffusion in the ownership of stocks, bonds, savings deposits, life insurance policies, and so on, among the wage-earners, must we not conclude that property and income are becoming more equally distributed? Not at all, since the middle groups and the rich are also accumulating more. But must we not at least conclude that with the continued increase in per capita income, capital accumulation would not be slowed down even though income were more equally distributed? The answer is not certain. There are signs that human wants and desires are increasing even faster than wealth and income, and it is by no means clear that it is easier now than formerly to save. If indeed the increase in average income should result in an increase in the per cent of the national income which is saved, a lower interest rate would tend to prevail and this in itself would tend toward greater equality in income distribution. Thus inequality of income and capital accumulation interact upon each other.

Functional Distribution. We turn now to a brief consideration of the statistical data relating to the distribution of income from the standpoint of the factors of production. According to Dr. W. I. King the national income was distributed in the four functional shares approximately as follows in 1910:

Wages and salaries	46.9 per cent
Interest	16.8 per cent
Rent	8.8 per cent
Profits	27.5 per cent

Dr. King, however, includes under profits the wage or salary that should be imputed to the individual entrepreneur — business man or farmer. He estimates roughly that three fourths of the profits so defined are really imputed wages and salaries. If we add these imputed items to wages and salaries we have the following distribution:

Wages and salaries	67.5 per cent
Interest	16.8 per cent
Rent	8.8 per cent
Profits	6.9 per cent

Similarly, the National Bureau of Economic Research estimated that in the ten-year period from 1909 to 1918 approximately 71 per cent of the net value product of mines, factories, and railroads went to wages and salaries, leaving about 29 per cent for profits, interest, and rent. As between salaries and wages, about 5.5 per cent of the net value product goes to salaries and about 65.5 per cent to wages.

With respect to the different industries the national income is distributed as follows:

	1913 (*In millions*)	1918 (*In millions*)
Agriculture	$5,887	$12,682
Mines, quarries, and oil wells	1,191	2,013
Factories	7,976	16,018
Construction	1,669	1,280
Other hand trades	852	1,704
Railroads	2,181	3,570
Other public utilities	1,154	1,661
Banking	509	767
Government (all branches of government in the United States)	1,829	5,352
Unclassified industries and miscellaneous income	12,332	15,318

Functional and Personal Distribution Contrasted. Functional distribution deals with the division of the social product among the four factors of production. It explains how the prices of the services of these factors are determined, but it does not explain the manner in which the earnings of these factors are distributed among the people who make up the community. For example, a business man may receive an income of $20,000 annually from his business, a composite of wages, interest, and pure profits. The theory of functional distribution explains why he receives that amount but it does not tell us what he will do with it. A part, we know, will be taken by taxation and used to support various governmental services in which the people of the community share according to their needs for these services. Another part will be consumed by the man and his family, and within the family group the distribution is in the main according to need. Still another part may be donated to the support of charitable institutions, whose benefits will also be shared according to the principle of need. The ultimate distribution of the annual real income of society is not made according to the principles of value and distribution, because those principles do not control the disposition of money incomes after these incomes have passed into the hands of the owners of the factors of production, but they do control the amount of money income the owners of these factors receive.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The number of stockholders in the United States has increased considerably, it is said, during the last fifteen years. Should you conclude from this that ownership of property is less highly concentrated than formerly? Why, or why not?

2. In the United States 3 per cent of the population receive about 22 per cent of the total income, whereas in the United Kingdom the same proportion of the population receive about 33 per cent of the income. How do you account for this difference?

3. Inequality in the distribution of income is not nearly so great as inequality in the distribution of property. Why is this the case?

4. Rank in order of importance the following causes of inequality in the distribution of wealth and income: inheritance of private property, unequal educational and social opportunities, differences in native capacity, chance gains, monopoly power.

5. "Let me emphasize again what I have said before, that it is probably in and through the exercise of the principle of plunder or the undue exercise of advantage, of gambling or of its allied principle of monopoly, or of special privilege or favor of some kind that many, very many, if not most of the greatest fortunes have been won." — J. W. JENKS. Evaluate this statement.

6. According to Professor Sorokin, in his study of "American Millionaires" (*Journal of Social Forces*, May, 1925), 39 per cent of deceased American millionaires started life poor and 30 per cent started life rich. Of living American millionaires only 20 per cent started life poor and 53 per cent started life rich. What does this difference indicate?

7. Professor Sorokin finds that of the wealthy men who started their career poor, 13 per cent had received no formal education, 58 per cent had an elementary education, 17 per cent a high-school education, and only 12 per cent a college education. Of the wealthy men who started life in moderate circumstances, 2 per cent had no education, 28 per cent had elementary education, 27 per cent high-school education, and 42 per cent college education. Of those who started life rich, 7 per cent had only an elementary education, 13 per cent had a high-school education, and 80 per cent a college education. All told, 54 per cent of the American millionaires held college degrees.

Of those who started their career poor, only 2 per cent became rich before the age of 30, 23 per cent between the ages of 31 and 40, 39 per cent from 41 to 50, and 36 per cent above 50; of those who started life rich, 27 per cent became responsible managers of big business before reaching the age of 30, and 53 per cent between the ages of 31 and 40.

8. The following is taken from the *Minneapolis Journal* for July 27, 1924: "In St. Paul, Probate Judge Howard Wheeler issued a statement showing that 60 to 70 per cent of all adults dying in Ramsey County last year left no more estate than enough to pay their funeral expenses; that although 2353 adults died, only 860 estates were probated; that of that number only 25 left more than $50,000, 90 from $10,000 to $50,000, 121 from $5000 to $10,000, and 20 left an estate of no definite value or consisting only of a cause of action for wrongful death."

9. The trade unions in the United States have, in the last few years, established over a score of labor banks, and more recently, a union life insurance company. What bearing do these facts have upon the accumulation of property and the investment of income by the wage-earning class?

CHAPTER XXXII

SOCIALISM

The Extent of the Socialist Movement. The fundamental purpose of this textbook is to explain as fully as possible within the limits of one volume the economic phenomena of our present social order. Yet we cannot overlook the fact that the economic order is a changing one. Nor is it possible to overlook the fact that a considerable proportion of the voters of all modern countries are seriously dissatisfied with the present order and desire a fundamental reconstruction of our economic system. Out of a total of 30,000,000 votes cast in the 1919 election in Germany about 45 per cent were socialist and communist votes. In the 1924 election the combined socialist and communist vote fell to about 35 per cent of the total; but in the election of 1928 it increased again to about 40 per cent of the total. In the 1924 election in Great Britain 5,552,000 votes out of a total of 16,564,000 were cast for the Labor party, which has a socialist program. Thus exactly one third of the British electorate voted the labor ticket. In France the socialists elected 106 deputies to the lower house in the May, 1924, election, and the communists elected 29. In the smaller countries of northern and western Europe about one third of the electorate vote the socialist ticket. Moreover, in Russia we have seen a stupendous experiment in communistic socialism. Not only do socialist representatives constitute about one third of all the representatives in the popular branch of the legislative bodies in northern and western Europe, but in nearly all these countries socialists have held important positions in the cabinets, and four countries — Sweden, Germany, Great Britain, and Denmark — have had socialist governments. Many municipal governments, moreover, are controlled by the socialists. In the United States socialism as a political movement is weak in comparison. There is but one socialist congressman at Washington, and in no case has a socialist occupied an important state or Federal office. A few cities have had socialist mayors, notably Milwaukee, Minneapolis, and Schenectady. In 1920, of a total presidential vote of 26,700,000 the socialist vote was only 920,000. In the 1924 election the socialists for the first time in many years refrained from putting up a presidential ticket and supported the independent candidacy of Senator La Follette.

Origin and Development of Socialism. This extensive and widespread movement is a newcomer among political parties. As a political movement it started in France under the leadership of Louis Blanc in the forties. For a short time the socialists shared in the provisional government established after the revolution of 1848. Fifteen years later a socialist political party was formed in Germany under the leadership of Lassalle. This united with another socialist group, followers of the great socialist Karl Marx, in 1875. This party, known as the Social Democratic party, made rapid strides in Germany, and by 1890 polled 20 per cent of the total vote cast. In 1912 this was increased to 35 per cent. During the war the party split into three groups, and the movement is now divided between the socialists and the communists. In England small socialist parties were organized in the eighties and nineties, but it was not until 1906 that the British Labor party, which soon adopted a socialist program, was definitely organized. Its growth was rapid. In the 1923 election it won 191 seats in Parliament. In France the socialists participated for a brief time in the provisional government of 1871, but the excesses of the commune régime and the subsequent banishment from the country of leading communards completely disrupted the movement. In the eighties five different socialist parties emerged, some moderate and others revolutionary. By 1900 kindred groups had merged sufficiently so that the number was reduced to two — one revolutionary in its views and the other moderate. These united in 1905 and formed the Unified Socialist party. The vote increased from less than 900,000 in 1906 to 1,400,000 in 1914.

Utopian Socialism. Modern political socialism was preceded by a nonpolitical group of socialists, frequently termed utopians. The outstanding leaders of this group were Robert Owen in England and Charles Fourier in France. Both wished to establish voluntary communities, each to consist of from two thousand to four thousand members. Owen founded, in fact, a colony in New Harmony, Indiana, in 1825, and scores of Fourieristic communities were established in the United States in the forties. None of these colonies consisted of more than a few hundred members. All of them failed in the end, but several survived for a number of years. Owen's community was communistic, all property being held in common, and the proceeds being equally shared by all. Fourier's communities were merely associative: the property of the community was owned by a stock company, the shares of which were subscribed for by such members as had capital, and these shareholders were to receive a certain percentage of the surplus production, the rest being distributed according to labor and ability; production and consumption were to be on an associative basis, but there was to be no equal distribution of income, nor was private property to be dispensed with.

Both Owen and Fourier made their appeal to all classes of society, but more especially to the powerful and the well-to-do. They believed that the existing order was all wrong and that this unfortunate condition was simply due to human ignorance. No one had heretofore set forth the true and ideal social order. They thought they had discovered such an order and believed that it was only necessary to inform people about it. Rational human beings would see the reasonableness and superiority of the new order, and all groups and classes would adopt it. Such, in brief, was the essential point of view of the utopian socialists.

Marxian Socialism. In 1848 Karl Marx and Frederick Engels wrote the Communist Manifesto, which presented in brief but brilliant fashion the point of view of what they conceived to be a "scientific socialism." The Manifesto did not set forth an ideal order which people ought to adopt. Instead it was an interpretation of the evolution of economic society. It sought to explain what was going on in the economic order and to predict on the basis of current tendencies what the future would bring. The central doctrine is the "materialist interpretation of history," as Engels called it. All social and political history is but a record of the conflict of economic classes. An economic class has common interests because its members have the same occupational and income status. But its interests come into conflict with the interests of other economic groups; thus, the economic interests of landlords are not the same as the economic interests of renters, nor are the interests of capitalists the same as those of wage-earners. Often when a thing benefits one group, by that very fact it injures another. Thus wherever society is divided into economic classes a conflict of interests develops, and the social and political movements to which this conflict gives rise constitute history. Now, according to Marx and Engels the alignment of economic classes is a product of the prevailing mode of production and exchange.[1] Thus it was the widening of the market and the development of the domestic system of manufacturing and finally large-scale production that created the capitalist class on the one hand and the wage-earning class on the other. In brief, changes in the technique of production create new economic classes, these class alignments give rise to a conflict of interests, and this conflict of interests gives rise to the social and political movements which make up historical events.

In the Communist Manifesto, Marx and Engels sketched the historical rise of the modern capitalist class. They pointed out how the rise of this new class inevitably resulted in the downfall of the old feudal monarchies. The rise of this class into a position of preëminence found expression politically in the establishment of modern political democracies. Representative government based upon the suffrages of the middle or capitalist classes wrested political power and special privileges from the landed nobility.

---

[1] See Alvin H. Hansen's article "The Technological Interpretation of History," in the *Quarterly Journal of Economics*, November, 1921.

But capitalism, according to Marx and Engels, cannot grow and develop without at the same time pushing a larger and larger portion of the population into the ranks of wage-earners. Large-scale production more and more eliminates the middle classes, they maintain. Thus the development of capitalism inevitably gives rise to a powerful working-class movement. The working class will finally rise to a position of preëminent power and will thus control not only the political state but also industry itself. Thus the capitalists will be ousted from control, all means of production will be socially owned, and government and industry will be controlled by and carried on in the interest of the working class.

Marx and Engels thus predicted that socialism was inevitable. This prediction was based on two fundamental postulates: first, that the middle class would inevitably be wiped out by the development of large-scale production; second, that the working class would inevitably sink into deeper and deeper misery by the very nature of the capitalist system. If these two propositions were indeed true, a revolutionary upheaval would no doubt be inevitable. Whether or not a socialist society would then follow would depend upon its workability. As a matter of fact, however, capitalism has hitherto not eliminated the middle class, nor has it resulted in the progressive deterioration of the working class. It is quite true that the pre-capitalist type of middle class is declining relatively. The *proportion* of our population who are individual proprietors is decreasing. Yet the total number of individual proprietors is not declining: farming still persists as an individual-proprietorship business, and there still continues to be an enormous number of small merchants, bankers, and even manufacturers. In a city like Minneapolis one cannot fail to notice the many small manufacturing plants which may be seen in all parts of the city. The old type of middle class thus continues to survive in large numbers. But, more than that, the capitalist system has developed a new type of middle class peculiar to itself. The great body of technical experts, managers and superintendents, scientists, and professional classes, without which the highly complex modern industrial society could not possibly operate constitute a new capitalist middle class.[1] Not only do they constitute a middle class from the standpoint of income and social status but also, in large part, from the standpoint of property. The property owned by the new middle classes is not of the individual-proprietorship type: it consists largely of stocks and bonds, mortgages, insurance policies, and savings accounts. Many recent researches indicate that the middle classes are not becoming appreciably poorer relative to the rest of the population either as to property or as to income. It is true that in Germany a considerable proportion of the middle class has been wiped out by the fearful price upheaval that they have experienced since the war, but this is in no sense a normal development of capitalism. With capitalism reconstructed on a firmer monetary basis, there can be little doubt that the middle class will rapidly recover in the new Germany.

Marx and Engels were quite right in assuming that the middle class constitutes the great bulwark of the present economic order; they were quite right in assuming that if this class were wiped out, capitalism would receive a body blow which it might not be able to survive; but so long as the middle class persists and finds its economic interests rooted in the present economic order, capitalism will be difficult to overthrow.

Nor has the development of capitalism resulted in the impoverishment of the working classes; on the contrary, real wages more than doubled from 1850 to 1900, and so far as the United States is concerned, real wages are higher today than ever before. Marx believed that under capitalism new machinery was being introduced so rapidly that it necessarily supplanted labor, creating a larger and larger reserve army of unemployed, thus progressively lowering wages. As a matter of fact, however, the increased output consequent upon the introduction of machinery gave rise to an increased demand for products; and the increased profits due to the cheaper machine processes stimulated the accumulation of capital and resulted in the building and operation of numerous new plants which reabsorbed the labor supply. Thus machinery has not in the end displaced labor, but has instead increased the marginal product of labor and thus increased real wages.

---

[1] See Alvin H. Hansen's article "Industrial Classes in the United States in 1920," in the *Journal of the American Statistical Association*, December, 1922.

But improved conditions do not in themselves prevent the labor movement from becoming socialistic. Labor is eager, as are all groups in society, to get all that it can for itself, and in so far as the working class becomes convinced that socialism will offer more than capitalism, it will naturally become socialistic. But at any rate a revolutionary overthrow of capitalism is not likely in a democratic society in which the conditions of the wage-earners are constantly improving. It was a violent overthrow of capitalism that Marx and Engels predicted.

Evolutionary Socialism. By the last decade of the nineteenth century many of the keener minds in the socialist movement had become thoroughly aware that the tendencies pointed out by Marx were not working out as predicted. Moreover, the socialist movement was gaining in political power and influence, and it was perforce being pushed into a practical program of social reform such as labor legislation, social insurance, and taxation reforms. Thus both the theory and the practice of socialism were becoming evolutionary. In the eighties the Fabian socialists in England challenged the doctrines of Marx and sponsored a gradual reconstruction of society, which, however, would lead toward the socialist goal. In 1894 the Austrian economist Böhm-Bawerk wrote a crushing criticism of Marx's theory of value and distribution. Since Böhm-Bawerk's attack no one has ever been able to reconstruct the Marxian economic system. Finally in 1899 Bernstein, a German socialist writer, openly admitted the collapse not only of the Marxian economic theories but also of Marx's predictions as to the tendencies of capitalism. In France, Jaurès, the outstanding figure of the French socialist movement before the war, stood for an evolutionary program of social reform. He said that socialism would come piecemeal, so gradually that no one would be able to say when we had passed from capitalism to socialism. Thus political socialism ceased to look forward to a final revolutionary overthrow of capitalism. It buckled down to a program of social reform which it was believed would finally lead to socialism.

Syndicalism. In the meantime, however, there developed a new revolutionary movement — syndicalism. It had its origin in France in the decade of the nineties. Its essential philosophy contained elements derived in part from anarchism, in part from socialism, and in part from trade-unionism. Unlike the socialist movement, it was antipolitical in its methods, and placed its reliance upon sporadic strikes, the general strike, and sabotage (a method which involves staying on the job but interfering by slovenly work or destructive acts with the profits of the employer, thus forcing him to come to terms). Unlike the current trade-union movement, however, it did not lay the emphasis on immediate ends, such as wages, hours, and conditions of work, but instead on revolutionary ends, a radical reconstruction of society. A great general strike was finally to bring about a collapse of the present system. In its place there was to arise a new society which was to consist of a general federation of *syndicats*, or trade unions. The local unions were to be federated into city federations; local unions in different parts of the country belonging to the same trade or industry were to be federated into national unions; and finally these national unions and the city federations were to be federated into a General Federation of Labor. This in part was the type of organization already existing in the French General Confederation of Labor. Thus the existing trade unions were to be the growing cells from which would spring the new social order. The structure was already developed, ready to assume control of the new order. Instead of a bureaucratic socialist state controlling all industry, each industry would be controlled and managed by the *syndicat* or *syndicats* covering that field of work. But the ownership would inhere in the local federations of *syndicats* in the case of local industries, and in the general federation in the case of national industries.

Syndicalism is thus similar to anarchism in that its social structure is a loose decentralized federation of more or less independent units. But the basic unit of syndicalism is the trade union. All society is thus organized from the standpoint of the producer. With anarchism, on the other hand, there would be federations (purely voluntary) for all sorts of purposes — federations for agricultural and industrial production, and federations for consumption that would make provision for food, fuel, electricity, and dwellings.

Guild Socialism. During the last decade guild socialism has developed in Great Britain. It is a compromise between syndicalism and political socialism. It aims to retain the present state, which would own all the means of production. But industrial guilds (including all manual workers, technicians, and managers) would manage the several industries; thus, the coal miners' guild would control the coal industry, the railway guild would control the railroads, and so on. Prices would be controlled by the state, and industry would be taxed by the state for its own maintenance and the provision for public services. The proceeds remaining after the payment of this tax might be distributed by the guild to its members on any basis it saw fit. Another school of guild socialists would substitute the "commune" for the state; this would be composed of representatives from civic organizations and from consumers' and producers' guilds.

Communism. The Russian revolution has brought to the front another socialist group, who call themselves communists to distinguish themselves from the moderate political socialists. (Before the Russian communists were established in power they were known as Bolsheviks, "bolshevik" meaning the majority; at the time when the name was given they were the majority in the Russian Social Democratic party.) They are extreme and revolutionary. They believe that socialism can be introduced at once and does not need to wait, even in such backward countries as Russia, for a complete development of capitalism. They believe that socialism cannot be established without disenfranchising employers and property-owners and introducing a "dictatorship of the proletariat," or propertyless masses. They believe in the expropriation of property by the state without compensation to private owners. They have no use for the democratic political state, but wish to establish in its place a "soviet republic." The soviet state consists of "soviets," or councils of workingmen in the urban districts and peasants in the rural districts, these local soviets in turn being represented in provincial and regional soviets, and these in turn in the all-Russian soviet. The Russian Communist party, in spite of the fact that it has a membership of only some 600,000, has succeeded in retaining control of the Soviet government since the autumn of 1917. This remarkable feat has been accomplished by the following methods: (1) voting by show of hands instead of by secret ballot, (2) disqualification of opposition candidates on some pretext or other, (3) repression of opponents' propaganda, (4) communist control of press and schools, (5) ousting of noncommunist trade-union officials and substitution of communists, (6) ousting of noncommunist officials of coöperative societies and placing communists in control of the central organization, (7) officering the army by communists, (8) opening military schools only to communists, (9) the Terror, by which internal foes of the government (particularly those in sympathy with foreign intervention) were ruthlessly suppressed.

The Russian communists immediately proceeded to nationalize the land, but the peasants were given "possession" of it for use, subject only to a levy on their surplus products. Next they proceeded to nationalize the factories, the mines, and the banks. By 1919 the process of nationalization was practically complete. But the extreme communism of this first period fell down on three counts: (1) it failed to maintain such capital equipment as existed, and it was utterly unable, except in sporadic cases, to make provision for new capital equipment; (2) it failed to secure the effective coöperation of the former owners, managers, and technical experts; (3) it failed to motivate the working class to maintain the former level of personal efficiency. Moreover, the communists were unable to import much machinery and railroad equipment from abroad, at first because they were engaged in a war with foreign countries on several fronts, and later because of the reluctance of foreign governments and capitalists to enter into trade relations with them. So long as the government was engaged in defending the revolution against foreign powers the peasant classes, constituting the overwhelming majority of the population, acquiesced measurably well in the communist rule; but as soon as the menace of foreign intervention disappeared the peasants became restless. Under the land law they were compelled to turn all their surplus products over to the state. This policy discouraged production on the one hand and resulted in secret barter on the other. The peasants demanded a repeal of the law conscripting all their surplus, and freedom of trade so that they could dispose of their products. This the communists were finally compelled to grant. Moreover, they yielded the right to lease land, to hire laborers, and to inherit a limited amount of property. Furthermore, the breakdown of factory production was so great that to get needed capital and increase production they were finally compelled to permit a partial return of private enterprise except in transportation, large industry, and export trade.[1] Concessions were also granted to foreign capitalists. Thus private enterprise and freedom of trade have to some extent again been introduced in spite of the fact that the communists still control the government. Private enterprise obtains in agriculture, in small-scale manufacturing and to a considerable extent in domestic trade; but the bulk of the large industries are organized in great state trusts. The new economic policy, with its partial abandonment of communism, has resulted in marked improvement in production. It is said that the real wages of Russian labor are equal to about one third that of American labor, but, on the other side, it is to be noted that an extensive system of social insurance has been introduced for the protection of the wage-earner. The relatively better condition of the urban wage-earners compared with the condition of the peasants has drawn a large number to the cities and created a serious unemployment problem. The prices of farm products have been inordinately low compared with the prices of manufactured goods and this maladjustment has created a difficult problem.

Socialism and Communism Contrasted. The democratic socialist parties of northern and western Europe which belong to the Socialist International are by no means in agreement with the Russian communists; they differ from them on the following specific points: (1) The socialists believe in political democracy, in universal suffrage, and in the rule of majorities; they are therefore opposed to the dictatorship of a small militant minority such as that established in Russia by the communists. (2) The socialists favor representative government which is based on geographical representation; the communists favor the soviet system of government. (3) The socialists do not believe that it is feasible to nationalize all industries at once; they believe that the public administration of industry is such a difficult and complex undertaking that it must be entered upon piecemeal. (4) The socialists do not believe in expropriation of private property without compensation. As an industry becomes nationalized or municipalized they expect to make full compensation to the private owners, perhaps through the issue of government or municipal bonds. In the end, however, private property in the form of investments will gradually disappear under the continuous pressure of inheritance and income taxes. (5) The socialists do not believe it is desirable to eliminate completely the freedom to engage in private enterprise. Only in the case of the great industries and municipal utilities, which are naturally monopolistic, do they intend to exclude private enterprise. They do, however, wish public enterprises or consumers' coöperative societies to engage in all fields and freely compete with private enterprise, and they believe that in such a situation private enterprise will in the end become relatively unimportant.

---

[1] The leading government industries are "transportation, textile, metal, mining, electrical, fuel, lumber, chemical and construction." See Harry W. Laidler's *A History of Socialist Thought*, p. 503.

There is much difference between socialists as to whether there should or should not be complete equality in the distribution of income. There are many moderate socialists who believe that it is necessary to have unequal rewards in order to secure the effective coöperation of persons with superior ability. In actual fact, since the socialists would undertake to socialize industry piecemeal, they would be compelled in the transition stage to adopt unequal rewards. Nevertheless the tendency in socialist theory is toward a system of equality.

Socialism and Capitalism. In view of the changing character of communism in Russia and the weakness of communist parties in the countries of western and northern Europe, the socialist program appears to be the only possible alternative to a capitalistic order. It is clear that the capitalist order is itself becoming more and more subject to social control, and that the distribution of wealth and income is subject to profound modification through the application of progressive income and inheritance taxes.

The changes in taxation and social control that have taken place in the last few decades go far to justify the description of the present order as a socialized capitalism. Yet it is still a capitalist order so long as private ownership and private management are the rule. Whether or not the function of government will continue to expand until the social order may properly be described as a socialist one, it is impossible for anyone to predict. But at any rate it is clear that there is a vast difference between a system of private property and private enterprise subject to extensive social control and a system in which public property and public enterprise is the rule. It is one thing to expect a certain amount of social control in private enterprise and property, and it is quite a different thing to undertake actively the public administration of all the hundreds of thousands of factories, shops, and farms. The program of socialization is by no means an easy one. Some of the most fundamental difficulties may be outlined for brief consideration.

Difficulties facing Socialism. 1. Will a socialist society be able to accumulate capital as efficiently as the present order? Will it be able to make adequate provision not only for the replacement of worn-out equipment but for additional capital? Without replacement the entire capital equipment of a country would be exhausted in a very few years. It is scarcely conceivable that any order could make better provision for capital accumulation than the present order; yet even under the present order the great masses of people save relatively little. Is it likely that the masses can effectively save collectively when they have not demonstrated their ability to do so individually? Socialists contend that under the new order incomes would be greater, and this would make possible more saving. Of course no one knows that incomes would in fact be greater, but even a very great increase in income is readily absorbed in a higher scale of expenditure without any appreciable increase in saving. Socialists point to corporate surpluses as an example of the manner in which capital will be saved under socialism, but it must be remembered that these corporate savings have been accumulated under the stimulus of private profit.

2. Will a socialist society, with its tendency toward equality in distribution, be able to motivate men with ability as effectively as does the existing order? The existing order offers large rewards for superior ability. To be sure, it would be very far from the truth to say that rewards are accurately proportionate to ability in the existing order. Moreover, it must be admitted that although under existing conditions able men demand high earnings, it does not necessarily follow that they would withhold their services if lower rewards were offered in a society where there were no better alternative opportunities. Leadership is to a great extent its own reward. Private enterprise, however, demands resourcefulness, initiative, and close attention to detail to a degree not likely to be equaled by a socialist order.

3. Will a socialist society be able to get as much work out of wage-earners generally as does the existing order? The history of self-governing workshops — a very extensive one both in England and the United States — is far from favorable on this point. Nor is the experience of municipal or national governments with their employees encouraging. The post-office service (if we except the numerous small sinecure offices) appears, however, to have attained a high grade of efficiency in the United States; but the postal employees do not control the administration, whereas under socialism the whole body of public employees would control the government. Would it be likely that such a government could impose hard and rapid work on its constituency?

4. Will a socialist society be able to secure as effective administration of industry as private enterprise? It is clear that there will necessarily arise a wide divergence of opinion among the members of a socialist society as to what might be done and how. There will always be a powerful minority trying to discredit the majority group and oust them from control. Now

a private corporation has this advantage: it can always carry out its policy in a single-minded and consistent fashion. If the minority stockholders do not like it, they have only the recourse of selling out and starting a new business of their own. Would not a socialist society always be hampered in its administration of industry by the squabbling of conflicting groups?

5. Will a socialist society be able to place all grades of workers in the various occupations as effectively from the standpoint of finding the fit job for the right man, and at the same time leave each individual as much freedom of choice in the matter of occupations as does the present order? A socialist society will certainly have majorities and minorities. Will not the majority seek to take all the more desirable places regardless of merit? To be sure, in the present order rich men's sons are often given desirable positions for which they are unfit; yet an organization that goes very far in that direction is apt to find itself beaten by a more vigorous and efficient competitor.

6. Will a socialist society be able to allocate the productive forces of society in the different channels of production more effectively than the present order? In other words, will it be able to substitute a better guide for production than market price, or might market price itself prove a safer guide for production under a socialist society, with more equal distribution of incomes? Market price is certainly not an *ideal* guide for production. The demands of not only the rich, but of all classes above the poverty line, turn a considerable amount of our energy toward the production of goods and services almost completely lacking in social utility, simply because they have the money and furnish the demand. There can be no doubt that market price can be an ideal guide for production only in a society which *utilizes* its incomes in a wise manner. But a society with an equal distribution of income may also conceivably utilize its income foolishly and wastefully. This is finally, then, a question of education. A wiser expenditure of income by all

classes in society is sadly needed. Nevertheless all modern societies are striving to attain greater equality in distribution, believing that a more effective utilization of income will follow for society as a whole.

Waste in the Present Order. On the other hand, we cannot be blind to the fact that the present order is very far from perfect. Not only does it present serious inequality in the distribution of wealth and income; in the field of production also grave deficiencies are not hard to find. Herbert Hoover, as Secretary of Commerce, has outlined succinctly the leading wastes of which our competitive order is guilty. His statement, in part, is as follows:

"I wish to make it clear that in speaking of waste, I do not mean waste in the sense of willful waste, but economic waste, which is the natural outgrowth of a competitive system. I do not mean the waste that any single individual can correct by his own initiative, but the waste that can only find remedy in collective action. Our industrial and distribution services are individually of the highest efficiency in the world. It is in their collective action that we can find progress.

"The kinds of waste that cause costly losses may be partly catalogued.

"1. Waste from speculation, relaxation of effort, and extravagance of booms, with the infinite waste from unemployment and bankruptcy, which comes with the inevitable slump.

"2. Waste from excessive seasonal character of production and distribution.

"3. Waste caused through lack of information as to national stocks of production and consumption with its attendant risk and speculation.

"4. Waste from lack of standards of quality and grades.

"5. Waste from unnecessary multiplication of terms, sizes, varieties.

"6. Waste from the lack of uniformity of business practices in terms and documents, with resultant misunderstandings, frauds, and disputes.

"7. Waste due to deterioration of commodities.

"8. Waste due to inadequate transportation and terminals, to inefficient loading and shipping, and unnecessary haulage.

"9. Waste due to disorderly marketing, particularly of perishables, with its attendant gluts and famines.

"10. Waste due to too many links in the distribution chain, and too many chains in the system.

"11. Waste due to bad credits.

"12. Waste due to destructive competition of people who are in fact exhausting their capital through little understanding of the fundamentals of business in which they are engaged.

"13. Waste due to enormous expenditure of effort and money in advertising and sales-promotion effort, without adequate basic information on which to base sales production.

"14. Waste due to unfair practices of a small minority.

"15. A multitude of wastes in use of materials, in unnecessary fire destruction, in traffic accidents and many other directions.

"These wastes are not the small change of industry and commerce. There is scarcely a step in this accomplishment of squeezing out waste which does not interpret itself in millions of dollars of annual saving.

"The work of the department during the last four years has demonstrated that there is a vast importance to these wastes. I am disposed to agree with a recent report of the engineering council that they amount in many lines to 25 or 30 per cent of the cost paid by the consumer or producer. They cannot all be corrected and where progress can be made it is only through toilsome building step by step in a thousand places, and always and only through the coöperation of well-disposed trade and industry, and through a wider understanding of the problems involved, and the coördination of effort necessary to secure results."

Value and Distribution in a Socialist Order. One other question remains to be considered. Should a socialist society ever be established, would there be any need to formulate principles relating to value, price, and distribution? Would such a society be one in which the products would be allocated by some central authority, or would it produce commodities for exchange at a value or price? Thoroughgoing state communism would presuppose the former, but socialism would not. If there is to be such a system as exchange and price, it is clear that the principles of price and value will be fully as important as in the present order. A socialist society would probably seek to establish a price corresponding to the cost of production. But to ascertain this cost of production would be as intricate a subject under a socialist order as under the present order. Certain socialists who adhere to the labor theory of value contend that the cost can be measured by the hours of labor required to produce the commodity. But it takes widely different numbers of hours to produce a bushel of wheat in Kansas, in Montana, and in New York. Economic rent thus emerges under a socialist order as truly as under the present order.

Nor is the matter in other respects as simple as the labor-value theorists indicate. An oak tree three hundred years old and a bushel of tomatoes may both be produced by the same number of hours of labor; but in the case of the oak tree the product does not emerge for three hundred years, whereas in the case of the tomatoes it emerges in a few weeks. One involves putting in labor for which there is no return for three hundred years. To put labor to work at something for which you can get no return for three hundred years is certainly a cost to a socialist society as well as to private individuals — a cost of waiting in addition to the cost of labor involved. The same argument will, of course, apply to all products made by elaborate machinery and plant. A socialist government could not place on a parity two products of which one requires a capital outlay of $5,000,000 and the other an outlay of $5000, even though the same labor (including the pro rata share of the labor cost of constructing the capital equipment) were required in both cases. Capital is scarce, and if one industry gets more, another gets less. Thus interest emerges in a socialist society, and the cost accountants will have to consider it in determining the prices of different goods. The conclusion is inevitable that principles of value applicable to the present order would also apply in large part to a socialist order.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. The English Wholesale Coöperative Society owns and operates a considerable number of manufacturing establishments. Its main business, however, is to supply goods at wholesale for the local consumers' coöperative stores. Suppose all manufacturing were brought under the control of the coöperative movement. Would interest be eliminated? profits? rent? Explain. Would a society with a universally developed coöperative system be more nearly like capitalism or like socialism? What points of similarity would it have to each? Would such a society have as great difficulty with respect to capital accumulation, efficient management, and discipline of labor as socialism would have? Explain.

2. "Higher real wages for the skilled labor, together with the increasing accumulation of property by wage-earners, will be likely to prevent the socialist parties from attaining a majority." Discuss.

3. "Should socialists actually gain majority control of Parliament in England for any considerable period, they are likely to find socialization of property ineffective. They are not likely to go far in this direction, but will then fall back upon the taxing power of the state. The net effect will be a socialization of income instead of a socialization of property." Do you agree? Why, or why not?

4. After the war the railroad unions sponsored the so-called Plumb Plan, which provided for the nationalization of the railroads and the leasing of them to a corporation managed by fifteen directors, five appointed by the president of the United States, five elected by the railroad managers, and five elected by the railroad workers. Which form of socialism does this plan resemble?

5. What is the fundamental distinction between the trade-union philosophy and the socialist philosophy?

6. Distinguish between the philosophy of the single-tax and the philosophy of socialism.

CHAPTER XXXIII

SOME ECONOMIC ASPECTS OF TAXATION

Government Finance and Economic Principles. The study of the financial operations of governments is called *public finance.* From the administrative point of view we may regard any governmental organization as a business undertaking. As such it bears many resemblances to a private business. Systems of accounting and control must be established, labor and commodities must be purchased and applied to the production of other goods and services, income and expenditure must be balanced, and the personnel of the working force must be organized and directed. Moreover, many of the important economic principles (the principle of diminishing returns, for example) which must be observed by the private business are also significant for the conduct of governmental business. On the other hand, there are important differences between private business and governmental business. Private business seeks the highest total profit as its goal; the government seeks the maximum public welfare. The income of private business is derived from the sale of commodities or services; the income of a governmental organization is chiefly derived from taxes imposed on private wealth and incomes. It is a relatively easy matter for the manager of private business to determine whether it is profitable to produce a given commodity or service, but it is not always possible to discover whether a given governmental service is worth what it costs the taxpayers.

From the social point of view governmental business, or, more briefly, "the government," may be looked on as a means for accomplishing certain ends which we have decided should be taken from the hands of private business, or for the provision of certain services which no private business would find it profitable to carry on. The rule of private business is that only those commodities and services will be produced for which the consumers are willing and able to pay a price at least sufficient to cover costs of production. But certain necessary public services, such as military protection and the maintenance of law and order, cannot be provided by a private concern, or at least they will not be provided in a satisfactory manner. Another rule of private business is that only those who are able and willing to pay the necessary price will be provided with the commodity or service. But there are many services, such as elementary education, sanitation, police protection, and the public highways, which we believe should be available to everyone regardless of his ability to pay; hence we turn these services over to governments, in whose hands they are provided for the use of everyone who may properly be said to require them, and their cost is met by taxation. There are exceptions to this rule (as we shall see later) when the users of the certain governmental services are required to pay all or a part of the costs of providing them.

When the government renders a service which is offered without charge to all who may require it and then raises the funds to defray its cost by a levy on the wealth and income of the citizens, it usually happens that the distribution of the wealth and income of the people is affected to a greater or less extent. Private incomes are very unequally distributed, and the same thing is true of private wealth. Many people who require large amounts of governmental service are unable to pay proportionate amounts of taxes; they receive proportionately more from the government than they contribute in taxes. Others, who own large amounts of wealth or receive large incomes, contribute proportionately more to the cost of maintaining the government than they receive from it in services. It must be noted that the difference between contribution and benefits received is relative. The receiver of a large income could not do without governmental service; his income would vanish were it not for the activities of the state; in a certain sense it is therefore correct to say that his entire income depends upon such governmental services as the maintenance of laws controlling the use of property. He does not contribute absolutely as much to the support of government as it is worth to him; but he contributes in most cases relatively more than the receiver of a small income, whose well-being is likewise completely dependent upon state activities. Since the effect of governmental services under our present system is to cause some to contribute relatively more than the relative benefits received by them and others to contribute less, we see that taxation has an important effect on the distribution of wealth and income.

The Nature of a Tax. A tax is defined as a compulsory payment collected by the government from individuals and corporations, without reference to benefit, for the support of governmental operations.

In the first place, a tax is a compulsory payment. Sometimes the government offers goods for sale in the market like any other seller. Those who wish may avail themselves of these goods by paying the price; those who do not wish to do so may escape payment. Such governmental charges are *prices* and not taxes.

In the second place, a tax is always collected for the support of governmental functions, the courts having consistently held that the power of taxation may not be employed to provide subsidies which merely enrich some individuals at the expense of the general body of taxpayers. The objects upon which tax money is expended must always be "public purposes." They must always be objects that contribute to the general welfare of the community. Under ordinary circumstances it would not be legal to levy a tax for the purpose of building a house for a particular person, since the service derivable from the house would benefit the person to whom it was given without benefiting the community at large.

Public services, such as good roads, education, and police, are paid for by taxation because they confer a general benefit on all the people. The gain of individuals from these services is either incidental to the public benefit or it is received by individuals as members of a community in common with all other persons in the community.

In the third place, a tax which has been legally imposed must be paid by the individual regardless of the benefit he may receive from its expenditure. A wealthy taxpayer cannot refuse to support the public schools because he has no children to send to them; nor can he get his taxes reduced on the ground that the benefit which the school system confers on him or his property is not proportionate to the amount of his taxes.

Nontax Revenue. A certain amount of revenue is raised by governments from sources other than taxation. These nontax sources may be classified as follows: (1) prices (that is, receipts from commercial enterprises), (2) fees, (3) special assessments, and (4) fines and penalties.

The term "prices" applies to revenues received by the government from the sale of commodities and services. The Federal government sells lumber from the national forests and power from its hydroelectric plants; city governments often own and operate water systems and charge the users according to the amount of water consumed. Sometimes these prices are essentially the same as the prices charged by private businesses. Lumber from national forests is sold in the competitive market for whatever price it will bring. In other cases the government monopolizes the service and charges prices that may be either above or below the prices which would prevail if there were competition. Like other monopolists the government may discriminate between different users.

A fee is a payment imposed on individuals by the government in respect to some service performed by it for the benefit of those individuals. It differs from a price in that the service for which it is charged is more important for the public at large than is the service for which a "price" is charged. The service is usually undertaken by the government for the purpose of regulation and control. Typical local fees are those charged for the inspection of building plans, boilers, and electrical wiring and for the recording of deeds to real property. The amount of the fee should not exceed the cost of rendering the service, and it may be less than the cost; in practice, however, it sometimes exceeds the cost of the service.

A special assessment is a kind of fee. It is charged when a city (it may be imposed by other governmental organizations) improves the streets or parks. It is paid by the owners of the property benefited by the improvement. Here again a public purpose is involved, and therefore it is proper to compel the owners of all taxable property to contribute. It would be possible to finance street improvements out of money raised by taxation; but since it is generally true that such improvements confer pecuniary benefit on the owners of adjacent real estate, a special charge is properly collected from them. In certain instances, of course, when the owner of the real estate is benefited little or not at all, all or a part of the cost of the improvement may be borne by the taxpayers at large.

Kinds of Taxes. There are many different kinds of taxes. Everyone is familiar with the local taxes on real estate and personal property and with the Federal income tax. These are all *direct* taxes. On the other hand, there are many taxes, such as customs duties and the duties on tobacco and other commodities, which are *indirect* taxes. Direct taxes are borne by the receivers of the incomes and the owners of the property upon which they are levied; indirect taxes are added to the prices of the commodities on which they are imposed and are borne by the consumers of these commodities.

Another classification of taxes distinguishes them according to the base upon which they are levied. The base of a tax is the taxable thing according to the quantity of which the amount of the tax is determined. For example, the base of the local tax on real property is commonly the sale value of the taxable property. If the sale value of the property is large the tax will be large; if it is small the tax will also be small. In the United States taxes are principally imposed according to the following bases: (1) the value of real estate, (2) the value of personal property, (3) net incomes, (4) gross earnings, (5) commodities, and (6) acts. All these bases are self-explanatory except possibly the last. Taxes on acts comprise the duties payable on certain legal transactions (such as the recording of wills and deeds), on powers of attorney, on the making of contracts, and the like. As sources of revenue they are unimportant, and the general opposition to them by the public indicates that they will remain so.

The Incidence of Taxes. When a tax is levied upon any kind of property or income it must be paid by the owner of the thing taxed. But the owner may be able to recover the tax from someone else who is a buyer or a seller of the taxable thing. When a person on whom rests the legal obligation to pay a tax passes it on to someone else, he is said to *shift* the tax. The person who finally pays, or the property from the income of which the tax is ultimately deducted, is said to "*bear*" the tax, and the *incidence* of the tax is said to be either on this person or on the property in respect to which the tax must be paid.

When a tax is imposed on producers' capital or on a particular commodity, it may have effects similar to those that follow from an added cost. For example, if a tax is imposed on a commodity and made payable by the producer, it may be passed on to the consumer, just as a rise in wages might be. But the notion that all taxes on business or on incomes from business are shifted to the consumer is erroneous. The first condition necessary for the shifting of a tax is that there shall be some untaxed field to which the capital, labor, and land subjected to the tax can withdraw. If the producers of cloth were taxed on each unit turned out, the only method by which they could pass the tax to the consumer would be by raising the price. This would cause a decline in sales, and unless the producers were prepared to accept a reduction of output they would not be able to shift the tax. A reduction in output can come in two ways: the factors may be withdrawn to some field where taxes are lighter, or the factories may run at less than capacity. If the former method is open to the producers, they will be able to shift the entire burden of the tax at once; if it is not, they may gain an advantage by curtailing production, raising the price, and waiting for the natural growth of demand to absorb the entire output of their factories again. The two methods are evidently the same, for the latter involves for its success the exclusion of new capital and the diversion of it to untaxed fields. Since taxes on commodities usually do not hit many sorts of goods, most of them can be shifted after a longer or shorter period of time.

It is said that a tax on economic rent cannot be shifted to the consumer of the products of the land which yields the rent; but if the tax is imposed on the rent of a tract of land only when it is used to grow wheat, and if at the same time the owners can derive approximately as great a return by growing some other crop, the tax will be shifted. The tax reduces the rent which the owners receive; and if this reduction is so great that they can obtain larger returns from their land by producing some other crop, they will produce that crop and not wheat. The supply of wheat having fallen off, its price will rise, and the consumer will bear the tax. But if rents from all the uses of the land are taxable, the tax does not affect the owner's choice of products, their relative supplies will remain as they were before the tax was imposed, and the tax will not shift.

In the second place, a tax which hits only surpluses cannot be shifted. By a surplus we mean a gain in excess of the expenses of production, including imputed expenses for the entrepreneur's land, capital, labor, and management. Fortuitous profits above the marginal cost of production, good and bad years taken together, may be taxed away without affecting the supply of the commodity. This statement does not mean, of course, that the state can take all gains from investments in excess of the riskless rate of interest: a part of these surpluses must be used to offset unavoidable losses.

A tax on net income usually cannot be shifted if it hits all occupations or industries alike. Since the tax does not increase or decrease with the amount of goods produced or with the amount of work done by the taxpayer, but with the net receipts from such economic activity, the individual can only escape it by causing his net income to decline. This no one would do as long as he had in mind the maximization of his income; for a hundred dollars received is better than fifty dollars, even though the government takes 10 or 20 per cent of either. The foregoing reasoning applies to corporations as well as to individuals. It is possible that an income tax at extremely high rates would deprive the taxpayers of so large a part of their incomes from labor and capital that industriousness and saving would be discouraged, and that as a result the social income would be reduced and the distribution of that income changed. But these effects would not prove that the tax had been shifted, because the tax would still be a deduction from the incomes of those on whom it was imposed.

When a tax hits margins it is more likely to shift than when nonmarginal incomes, firms, etc. are hit. A business tax that impinges on marginal producers in any field will drive them out of that field if there is any other to which they can go and escape the tax. But even though the tax is not general, it is less likely to be passed on to consumers if it is imposed on firms that are making a pure profit. These firms receive for their goods a price that more than covers all expenses (including imputed expenses for proprietors' capital), and they will be less likely to desert the taxed trade when compelled to pay a tax than would the marginal firm.

A tax on monopoly revenue cannot be shifted. By this we mean that if the tax subtracts a certain percentage from the excess above cost of production under competitive conditions which a monopolist receives by virtue of exclusive control of supply, the monopolist cannot compel consumers to reimburse him. By assumption the monopolist is already receiving the price which will yield the maximum profit; and if he raises the price, his maximum net income will be reduced and he will still have to pay the tax at a certain percentage rate on the monopoly revenue. On the other hand, if the tax is imposed on the commodity produced by the monopolist and if he is required to pay it, the tax will probably shift, because the monopolist, like any other producer, can withdraw his capital and reduce the output. He can thus escape part of the tax and at the same time raise the price to the consumer. The capital withdrawn can be invested at the competitive rate free from this special commodity tax.

The details of the process of shifting are very complex. Here we have only attempted to set forth the most general principles. A tax is not easily shifted (1) when it is universal — when it hits all people and all businesses alike; (2) when it hits net income; (3) when it hits surpluses and does not impinge on margins.

In the preceding discussion we have assumed that the agents of production are completely mobile. This assumption must now be qualified. If a commodity tax is imposed on producers they may bear it for a time because they are unable to change the supplies of goods which they are bringing to market. Their capital may be highly specialized, or their land not easily turned to other uses. But all capital wears out in the long run, of course, and when the wearing-out process has begun to reduce the supply of the commodity the tax will begin to shift. Even the labor of the proprietors may not be readily adaptable to other types of production. The supply of the product being fixed, there is no method by which they can force the tax on the consumer.

The ease with which a tax on commodities can be shifted to the consumer is dependent upon the elasticity of consumers' demand for the product. If the demand is very elastic, it may require a large reduction of supply to raise the price by the amount of the tax; if it is very inelastic, only a small reduction will be necessary.

A tax may be shifted backward as well as forward; for example, if a heavy commodity tax is imposed on the producers of a given good, they will try to pass it to the consumers by reducing output. But a reduction of output not only curtails the amount of the finished commodity; it also curtails the demand for labor and raw materials. If the labor used in making the commodity can easily move to other occupations, and if the land and capital with which the raw materials are produced can easily be turned to other uses, then the consumer will have to pay the tax; but if the labor and the land are not industrially mobile, then either or both of them may be compelled to submit to lower returns and thus be compelled to bear part of the burden of the tax.

Finally, we must distinguish between the shifting of a tax and its *effects* on property values and individual wealth. If a very heavy tax were imposed on grocers it would, under ordinary conditions, be shifted, because some of the people engaged in the grocery business would no longer be able to meet their costs and would therefore be forced to withdraw. In the process of withdrawal many would lose money. They would be forced to sell their fixtures at a sacrifice, to abandon such patronage as they had built up, and to seek different occupations at lower wages. But the tax might still be shifted in its entirety to the consumer. Those who remained would be able to cover their costs, including the tax, and the prices of goods would rise enough to cover the tax as well as other expenses. The loss sustained by the owners of abandoned grocery stores would be an effect of the tax; its incidence would still be on the users of the goods.

The Capitalization of Taxes. When a tax hits a durable good, either in its primary incidence or ultimately, it tends to be capitalized. This phenomenon can be illustrated best in the case of land. Let us assume that a parcel of land yields $100 rent annually and that the rate of interest is 5 per cent. The land will then be worth $2000. Now let a tax of $20 a year be imposed on the rent. The rent which any owner can count as net will be reduced to $80 and the value of the land to $1600. To put the same calculation in another way: the capital value of the land before the tax was levied was $2000; the capital value of the tax ($20 a year at 5 per cent) is $400; this amount is subtracted from the original sale value of the land, leaving a new capital value of $1600.

The first condition necessary for capitalization has already been stated: the thing taxed must be durable. If it is not, if supplies are constantly coming on the market, these supplies will be checked when the tax is imposed, and the burden of the tax will be passed forward to the consumer; if the durability of the good is considerable but not comparable to land or to land improvements, then there may be a partial capitalization of the tax. The second condition is inequality. It is often said that taxes on rent are always capitalized, but this conclusion is hardly justified by theoretical reasoning. The problem turns on the question whether after the tax has been imposed the first purchaser of the land will be able to compel the seller to accept a reduction in its price equal to the capitalized value of the tax. The investor in land, like any other buyer of capital goods, seeks to put his money where, other things being equal, it will earn the maximum return. If he finds that land has been subjected to a new tax and that other permanent producers' goods have not, he will be able to force the original landowner to accept the loss of $400 by threatening to buy other property not affected by the tax. Concretely, he will say to the landowner that he can invest $2000 in machinery, in buildings, or in merchandise and receive $100 annually free from any tax deduction; and if all prospective purchasers of land have the same opportunity the landowner will be compelled to accept $1600 if he wishes to sell. On the other hand, if the prospective purchaser finds that all alternative investments are affected by the tax to the same extent as land is, he will be unable to compel the landowner to accept $1600, and the tax will not be capitalized.

The Amount of Taxation. The amount of taxation imposed by any government is determined in large part by the amount of its expenditures, but not all the money receipts of any government in a given year are derived from taxes. Part ordinarily comes from fees, part from public enterprises such as waterworks and the post office, and part from loans. That the first two constitute independent sources of revenue is easily understood. Loans, on the other hand, appear to be mere anticipations of revenue from taxation or from nontax receipts. This is usually the case when borrowing takes place for the purpose of financing permanent improvements such as public buildings, roads, sewers, and waterworks: the government borrows to obtain immediate funds, and the bonds which it issues are later retired by payments from the charges made to consumers or by funds raised by taxation. When loans are made to meet the expenditures arising from emergencies such as wars, governments frequently attempt to wipe out at least part of the indebtedness by taxation when the emergency is past. But they do not always do so; hence borrowing for these purposes may cause an increase in receipts which is not offset later by taxation. Moreover, although the original indebtedness for permanent improvements may be wiped out, the time comes when the improvement must be replaced and further loans are made for that purpose. Public debts may thus become permanent. Yet in spite of these discrepancies the amount of public expenditures is a good measure of the amount of taxation required in any community.

The control of the amount of taxation depends primarily, then, upon the control of the amount of public expenditures; but the determination of the legislature to spend is partly controlled by the amount of taxation which it is thought advisable to impose. Hence we conclude that the volume of public expenditures and the volume of taxation mutually govern each other: if it is easy to raise large amounts of taxation, public expenditures are likely to be large; if there is great need for public expenditures, the amount of taxation is likely to be great. On the one hand, the ability of the people to sacrifice a portion of their incomes for the support of public services must be considered (a population whose income is so small that it barely affords the necessities of life cannot contribute large amounts to the support of governmental services); on the other hand, the utility that can be obtained from the expenditure of public revenue is also a determining condition. Consider the case of public highways. Before the automobile became common, graveled or dirt roads were generally thought to be adequate for rural districts; but with the growth in the use of motor cars the demand for highways with a continuously smooth and firm surface became insistent. Money was borrowed and taxes were imposed to raise the funds for improving them. In this case the additional utility derivable from better highways dictated the increase in taxation.

But the need for improvements does not alone control. Some communities are too poor to support good roads, good schools, and expensive public buildings; hence, notwithstanding the utility of these goods and services, they are not provided, and taxation remains relatively low in absolute amount.

A third factor that enters into the determination of the amount of taxation is the efficiency of government. If the administrative departments of government are efficient, people are more likely to consent to the imposition of heavy taxes than if they are corrupt and inefficient; yet we find that taxes are sometimes very high in states and cities where the standards of efficiency and honesty among public officials are very low. This is because there are still other forces at work to determine the amount of taxes. Fourth, heavier taxes can probably be imposed if the tax system includes many indirect taxes than if it is composed chiefly of direct taxes. The reason is obvious: indirect taxes are shifted to the users or consumers of the articles on which the taxes are first imposed; they are not paid ultimately by those who must first advance them.

The ultimate payers are often not cognizant of the fact that they are paying a tax when they buy goods or ride on street cars, and when they are aware of it the exact amount of their payments may be unknown to them. A user of tobacco will ordinarily pay more willingly a tax of $10 when it is added to the price of his cigars or cigarettes than he would if it were assessed against his income or his real estate.

The attempt has often been made to lay down rules for determining the proper amount of taxation to impose in any given community. As we have seen, the amount depends upon the available income from which taxes can be paid and upon the value of the services which the public receives from public expenditures. This much is generally admitted. But the foregoing statement is indefinite. It does not tell us whether it is good economy to take 5 per cent or 20 per cent of the income of the people for the support of governmental services. The impossibility of laying down definite statements is inherent in the nature of the problem. The conditions that have been stated in the earlier portions of this section are not constant. The valuation which taxpayers set upon public services varies with the changing conditions of industrial life, with changes in the opinions of the voters as to the proper degree of state interference, and with changes in the efficiency and technique of public services. The income available for taxation is also variable, changing not only with the variations in the total income but also with changes in its distribution.

Increase in Taxation. Any change in the amount of public expenditures is certain to affect the amount of taxation. And for more than a century the amount of public expenditures has been persistently on the increase. In the United States this increase has been relatively slow in times of peace; but two great wars, the Civil War and the recent World War, have been the occasions if not the causes of sharp upward trends in taxation. That a great war should cause an increase in the expenditures and taxation of the national government is easily understood, but that it should also be followed by a rise in local taxation is not easily explained; nevertheless the evidence is clear.

TABLE XXXVI. VOLUME OF TAXATION IN THE UNITED STATES FOR 1890–1926 (IN MILLIONS OF DOLLARS)[1]

IN "CURRENT DOLLARS"

TAXING AUTHORITY	1890	1903	1913	1919	1921	1922	1923	1924	1925	1926
Federal	$374	$521	$668	$4500	$4905	$3487	$3032	$3193	$2966	$3207
State	96	155	307	570	783	858	917	1017	1107	1264
Local	405	706	1219	2395	3150	3157	3285	3611	3818	4084
Total	$875	$1382	$2194	$7465	$8838	$7502	$7234	$7821	$7891	$8555

IN "1913 DOLLARS"

TAXING AUTHORITY	1890	1903	1913	1919	1921	1922	1923	1924	1925	1926
Federal	$465	$609	$ 668	$2180	$3339	$2343	$1972	$2133	$1869	$2124
State	119	181	307	276	533	577	597	679	698	837
Local	503	826	1219	1160	2144	2122	2137	2412	2405	2704
Total	$1087	$1616	$2194	$3616	$6016	$5042	$4706	$5224	$4972	$5665

The upper division of Table XXXVI shows that the increase in the amount of taxes collected was greater between 1913 and 1926, a period of thirteen years, than it had been in the preceding twenty-three years; but since the purchasing power of the dollar was increasing during the greater part of the decade 1890–1900, and since the decline in its purchasing power was relatively slow before 1914 as compared with the period of the war, a comparison of the absolute amounts does not accurately measure the actual amount of goods represented by the taxes. In the second division of the table the data have been reduced to a common basis. From these figures we discover that in 1926 the collections of the Federal government were 457 per cent of the collections in 1890 and 318 per cent of those in 1913; for the states the collections in 1926 were 703 per cent of the collections in 1890 and 273 per cent of those in 1913;

---

[1] From *Cost of Government in the United States, 1925–1926*, pp. 70, 76. By permission of the National Industrial Conference Board.

for local governments the percentages were 538 and 222 respectively. The total collections of all three types of governmental organization in 1926 was 521 per cent of the collections in 1890 and 258 per cent of those in 1913. Since 1921 the Federal taxation has declined, though not to a marked extent, while the taxation of the state and local governments has increased.

The increase in the amount of taxation does not measure the burden of taxation. During the entire period covered by Table XXXVI national income was increasing and a larger amount was available for paying taxes. The burden of taxation depends upon the proportion of the income taken by taxes and not on their absolute amount. In Table XXXVII the amounts of taxation raised by the three governmental divisions are compared with the national income.

TABLE XXXVII. TAXES AND THE NATIONAL INCOME, 1890–1926[1]

YEAR	NATIONAL INCOME (MILLIONS)	PERCENTAGES THAT TAXES ARE OF THE NATIONAL INCOME			
		Total	Federal	State	Local
1890	$12,080	7.2	3.1	0.8	3.3
1903	20,500	6.7	2.5	0.8	3.4
1913	34,400	6.4	1.9	0.9	3.6
1919	68,260	10.9	6.6	0.8	3.5
1921	55,597	15.9	8.8	1.4	5.7
1922	61,633	12.2	5.7	1.4	5.1
1923	71,558	10.1	4.2	1.3	4.6
1924	70,768	11.1	4.5	1.5	5.1
1925	77,313	10.2	3.8	1.4	5.0
1926	78,649	10.9	4.1	1.6	5.2

It is shown by Table XXXVII that from 1890 to 1913 the taxation of the national government declined with relation to the national income, that state and local taxation increased slightly, and that the total of all taxation declined. Immediately after the close of the war the percentage of the national income taken by the national government rose to about

---

[1] From *Cost of Government in the United States, 1925–1926,* p. 77. By permission of the National Industrial Conference Board.

9 per cent, and since the state and local governments also increased their collections the burden of taxation amounted to nearly 16 per cent of the total income. Since 1923 the percentage of national income taken by taxation has remained relatively constant at 10 to 11 per cent, although in more recent years there appears to be a slight tendency for the burden of state and local taxes to increase.

It must be emphasized that the percentages of the national income taken by taxation as shown in Table XXXVII depend both for their magnitude and for their trend upon the amount of the national income. There is no census of income and no other source from which the total can be accurately ascertained. The figures given are estimates, and, as is usually true of estimates, authorities are not entirely in agreement. The figures given for income are conservative by comparison with the computations of other competent statisticians;[1] yet even on the basis of these conservative figures the burden of taxation in this country is evidently not crushing. Moreover, it must be remembered that the ability to spare private income for the support of public services increases more rapidly than the national income.

Nevertheless the fact remains that governmental taxation and expenditure are increasing. The causes are not the same

---

[1] Below are given the estimates of the total current income of the people of the United States, published by the National Bureau of Economic Research in their *News-Bulletin* for Feb. 21, 1927, and the percentages of such income taken by taxation in various years.

YEAR	NATIONAL INCOME (IN MILLIONS)	TOTAL TAXATION (IN MILLIONS)	TAXATION AS PERCENTAGE OF INCOME
1913	$32,000	$2194	6.9
1919	67,254	7465	11.1
1921	62,736	8838	14.1
1922	65,567 [2]	7502	11.4
1923	76,769 [2]	7234	9.4
1924	79,365 [2]	7821	9.9
1925	86,461 [2]	7891	9.1
1926	89,682 [2]	8555	9.5

[2] Preliminary estimates.

for all three forms of government. The principal object of Federal expenditure is and always has been payment for past wars and preparation for future wars; for the state and local governments the principal objects requiring taxation are education and highways, followed in order of importance by protection to persons and property and by social relief.

Since the return of the Federal government to a peace-time basis, from 40 to 50 per cent of its expenditures have been for repayment of, and interest on, the national debt, between $1,617,000,000 and $2,039,000,000 having been paid out for these two purposes annually between 1923 and 1926. Since the debt is held at home these disbursements are not a "burden" in the same sense that expenditures for general government and the support of the army and navy are burdens. The money to make payments on debt is collected by taxation from the citizens of the United States, who also hold the bonds. The entire transaction is a transfer of wealth from the taxpayers to the bondholders, which constitutes no present burden upon the total income of the nation.

The increase in public expenditures is a phenomenon which has manifested itself for more than a century. It is not confined to certain countries: in different countries somewhat different objects of expenditure are responsible; yet it is generally conceded to be due (1) to the increasing cost of armaments, (2) to technological progress, and (3) to the growth of democracy. The progress in technology makes possible services which could not have been provided at an earlier time; the progress of democracy gives rise to a demand for provision at public expense of better services and of new services that progress in technology makes possible. In a democracy the mass of the voters, who on the average own very little property and receive relatively low incomes, have the power to vote expenditures and levy taxes. And since direct taxes bear most heavily on the receivers of large incomes and the owners of property, there is constant pressure for the provision of additional services at public expense.

Tests of a Good Tax System. There are certain qualities which a tax system should possess to make it administratively successful. These qualities or characteristics are essential and must be taken into account in any discussion of the effect of taxation on the distribution of wealth and income or on the prices of goods. In the first place, a tax system must be productive; it must yield adequate revenue before any other test can be taken into account. Second, the system as a whole should be flexible enough to permit the fiscal authorities to adjust the revenue to changing needs. Third, the taxes that make up the system should be dependable; that is, they should be so imposed that the revenues from them can be forecast with some degree of certainty. This is desirable in order that proper adjustment of receipts to expenditures may be made in advance of the collection of the taxes. Fourth, every tax ought to be economical. The costs of assessment and collection ought not to consume a large portion of the revenue that it yields; neither ought the tax to be imposed in such a manner that the taxpayers are greatly inconvenienced or put to much expense to comply with the rules and regulations governing its payment.

In addition to these administrative tests a tax or a system of taxes ought to meet certain nonadministrative requirements. It ought to be imposed in such a manner that the amounts of tax paid by people in different income classes conform to the sense of justice of the community in which it is levied, and it ought to conform to the political institutions of the government imposing it.

The Distribution of Taxation. Two principal theories of the distribution of the burdens of taxation have been advanced by writers on this subject. The benefit theory asserts that people should pay according to the benefits they derive from the operations of the government. An obvious but superficial argument in favor of this theory is that taxes are burdensome and ought therefore to be paid by those who occasion them. It is not the purpose of taxation, it is said, to favor some at the expense of others, and to avoid this it is necessary to tax people according to the governmental expense they occasion. The first objection to the benefit theory is its lack of practicability. Much of the activity of governments is made necessary by people who are too poor to contribute their full share to taxation. One can hardly maintain, for example, that unless a person of small income can pay the cost of educating his children they should not attend the public schools, or that unless he pays his proportion of the cost of police the state should not protect him. The theory is also impracticable because it is impossible to discover the individual benefits resulting from many government services. A military establishment maintained by the national government is a necessity, but it is impossible to make even a rough estimate of the benefit received from it by any given person. A second and more fundamental objection to the theory is the fact that it is based on an erroneous concept of the purpose of government. Organized society does not exist for the benefit of individuals as such, but for the welfare of the entire community. Governmental services are not designed to advance the interests of individuals except in so far as their advancement is an advantage to the community.

There is, however, a certain amount of admitted justice in the theory. One could hardly defend a tax on the people of Maine to build roads in California or a tax on the people of one city to support a fire department in another remote municipality. Or, again, since the owners of automobiles occasion the major expense of permanent highways it is only fair that they should pay special taxes for the support of this service. On the other hand, there are many who believe that the national government should assist in the maintenance of a system of education, and it is common practice for a state to tax all the people within its jurisdiction for the partial support of the schools in communities where, without this aid, passable educational standards could not be maintained.

A second theory asserts that taxes should be paid in accordance with ability to pay. This means in its most general form that all the people living within the jursidiction of a government should contribute to its support according to their ability as measured in terms of wealth or income. In support of this theory it is said that government is a common enterprise undertaken for the good of all. All are obliged to support it to the limit of their capacities if the need arises. In extreme cases, when public danger is great, the state may take any portion of a citizen's property needed, may circumscribe his freedom of action, and may even force him to enter the army and risk his life. Ordinarily so great a contribution is not required, but the same principle holds for the smaller need as for the greater one: taxes should be apportioned according to ability.

How ability to pay should be measured is not an easy problem to solve. Wealth is not an entirely satisfactory criterion, because we know that many people have incomes and therefore ability to pay without possessing much wealth. Nor do the capital values of different forms of wealth vary exactly with the income derived from them. Money income is also deficient as a test of ability. To begin with, some incomes are derived from the ownership of property; others are the results either of labor or of the active management of business. The former are commonly called *unearned* incomes; the latter are called *earned* incomes. To tax two individual incomes of the same size, one of which is earned and the other unearned, at the same rate is clearly not just. In the second place, some money incomes are certain and others are uncertain; some are temporary or limited in duration and others are relatively permanent. Incomes from labor and from the management of business are both more uncertain and more temporary than are incomes from property. Third, individuals having similar incomes may have greatly different obligations. An unmarried individual who receives $5000 has greater ability to pay than another who receives the same money income but who must provide for a family. These differences in ability are recognized in income-tax laws by allowing exemptions for dependents and by imposing lower rates on earned than on unearned incomes.

Still another theory, the equal-sacrifice theory, which is essentially a variant of the theory of ability to pay, asserts that taxes should be so imposed as to cause each taxpayer the same sacrifice. Here the burden of taxation is measured in terms of satisfaction that must be given up through the payment of taxes. The measure of the tax burden is conceived to be subjective. One obvious obstacle to the application of this theory in practice is the difficulty in ascertaining the amount of subjective loss occasioned by a given tax payment. It is, perhaps, not difficult to convince ourselves that to take 10 per cent of the income of a mechanic who earns $1500 a year imposes a greater sacrifice than to take the same percentage from the income of a lawyer who earns an income of $10,000 a year; but when the distinction must be made between individuals receiving more nearly the same incomes the application of the principle is not so easy. Moreover, we cannot be sure that to take the same percentage of income from two individuals who have the same incomes and the same needs will impose the same sacrifice, although there could be little injustice in assuming that it would. In conclusion, it may be admitted that the theory is workable in a rough way, but a rigid application of it cannot be defended.

None of these theories is broad enough to include all the considerations that should enter into the determination of the distribution of taxation. In some cases the principle of benefit is applicable. These are chiefly instances in which the benefit is assignable to a group of people living within a city or some similar political division. Usually, too, the benefit theory is applicable when a particular group receives a direct pecuniary return from a certain public service which is important to them as well as to the community at large. Hence we expect the people within a given political division to support their local government. But when the service to be supported is of great social importance and when the benefits cannot be assigned, we must resort to the theory of ability to pay or the theory of equal sacrifice to guide us in the distribution of the tax burden.

Some taxes which have the stamp of long approval on them cannot be justified on any of the theories we have set forth. A good example is the sumptuary tax, which is levied wholly or in part to discourage the use of a luxury or of a good considered detrimental to public welfare. Before the passage of the Eighteenth Amendment to the Federal Constitution the taxes imposed on alcoholic beverages, which were extremely heavy as compared with other commodity taxes, were a case in point. Again, neither benefit nor ability to pay can be used to justify the heavy import duties imposed on some commodities under the policy of protection. In most communities there is an opinion, often not definitely formulated, that landowners ought to bear heavier tax burdens than the owners of other sorts of property, because the rent of land is looked upon as an unearned income. The theory of benefit enters also, because the landowner is thought to receive greater gains from the growth of the community and from public improvements than other property-owners.

In conclusion, it appears that there is no one rule that can be rigidly followed in the apportionment of taxation. No rule that is not as wide as general social welfare will cover all cases.

Progressive Taxation versus Proportional Taxation. A progressive tax is one whose rate increases as the base upon which it is imposed increases in amount; a proportional tax is one whose rate remains constant. In the case of the progressive income tax the amount of the individual's net income is the base. As the amount of that income grows larger the rate of taxation is raised. Progression should not be confused with the exemption of small amounts of income or property from taxation. The income tax is not imposed on very small individual incomes, nor the inheritance tax on small estates. A small amount of personal property is usually exempted from the personal-property tax. The purpose of this exemption is twofold: in the first place, to collect small tax bills costs more than they bring into the treasury; in the second place, the exemption may be a concession to the opinion that the poor should not be compelled to contribute to direct taxes. But such exemptions have much the same effect as progression.

Progressive taxation must not be confused with the classification of taxable property. It is common for the states to tax some kinds of business at a higher rate than other kinds, or to levy one sort of tax on one or a few types of enterprise and a different sort on the remainder. Railroads and public utilities are often taxed on gross receipts while manufacturing companies, retail stores, farms, and other businesses are subject to the general ad valorem property tax.

Progressive income taxes are now imposed by the United States, Great Britain, and other countries. In the United States the introduction of a Federal income tax was made possible by the Sixteenth Amendment, which gave Congress power to impose income taxes without apportionment among the states according to population. Before the passage of this amendment and the law which gave it effect a long dispute was waged over the question of progression.[1]

---

[1] Under the Act of 1928 the Federal tax on individual net income is levied at the following rates:

Credits against net income for normal tax	
1. Married person, head of a family or husband and wife	$3500
2. Single person	$1500
3. For each minor child or other dependent	$400
4. Dividends from corporations subject to the income tax	
Normal rates	
1. First $4000 of net income above credits	1½ per cent
2. Next $4000 of net income above credits	3 per cent
3. Remainder of net income above credits	5 per cent
Surtax	
Income over $10,000 to $14,000	1 per cent
Income over $14,000 to $16,000	2 per cent

and so on up to $24,000, where the surtax is 6 per cent on the amount of income between $22,000 and $24,000. For incomes above the latter figure the rate rises by varying intervals to 20 per cent on all income in excess of $100,000.

Deduction for earned income

If the income is earned, the taxpayer is allowed to deduct 25 per cent of the amount of tax due on any amount of income up to $30,000. The first $5000 of income is assumed to be earned in all cases.

One of the first arguments for progression rises from the theory of equality of sacrifice, which we have already noticed in connection with the question of the distribution of the tax burden. That taxes may be imposed so as to cause each person to give up the same amount of satisfaction as every other, they must, it is said, be progressive. As income increases, its marginal utility decreases in accordance with the principle of diminishing utility; conversely, when income is decreased by taxation, its marginal significance rises. But persons who have large incomes derive a relatively small satisfaction from any given unit of income, whereas those of more moderate means derive a larger satisfaction from each unit of their incomes. If the tax is designed to cause all persons to contribute equally, it must deprive each of the same amount of satisfaction. But to do this it must take proportionately more from the rich than from those whose incomes are smaller. This argument for progression is subject to several objections. In the first place, one must agree that justice in taxation means equality of sacrifice before he can accept it. In the second place, while it may be granted that a tax of 10 per cent imposed on rich and poor alike deprives the latter of greater satisfactions than it does the former, it does not follow that the same distinction exists when the same rate of taxation is imposed on two rich men who have different incomes. Concretely, it does not follow that to take $3000 from one who would otherwise receive $30,000 a year deprives him of any greater sum of utilities than to take $10,000 from another who would receive $100,000. That is to say, we may well doubt whether the principle of diminishing utility applies with sufficient accuracy to money income above an amount which will purchase all reasonable luxuries of life to justify a system of progression. A final objection to the sacrifice theory as an argument for progression is its lack of practicability. Even if it is granted that units of income become less and less significant as total income increases in amount, there is no method of ascertaining the rate at which this decline takes place.

A second and more practical reason for a limited amount of progression is the fact that most import duties and internal taxes on commodities bear more heavily on small incomes than on large incomes, because those who receive the larger incomes do not consume the taxed commodities in proportion. Such taxes are said to be regressive, and it is argued that they should be offset by some other tax that is progressive. The argument is good, though it is not capable of exact application, since we do not know just how great the regressivity of the other taxes is. And granting that the wage-earners and small proprietors do pay a disproportionate share of the commodity taxes imposed by the national government, we must still take into account the relatively heavier share of the large property-owners in state and local property taxes.

It is also urged that large fortunes are very frequently acquired at the expense of the public, and progression should be used to recapture a portion of the wealth thus unjustly obtained. The argument goes on to show that the progressive income or inheritance tax would aid the public to recover wealth which has been taken from them by unsocial acts. This is clearly not an argument of much weight. It is impossible to determine how many of the very large incomes in this country are predominantly the result of predation. Certainly the state should not attack *all* wealth until it has good evidence that at least a large majority of the big fortunes have been acquired by unsocial methods of making money.

Progression has been applied in the United States to both income and inheritance taxes, and there appears to be scant probability that we shall abandon it in the near future. The reasons for its apparent permanence are (1) the theory of equality of sacrifice, (2) its superior productiveness as compared with any proportional tax that is politically feasible, (3) the regressiveness of commodity taxes, and (4) the desire of many persons to bring about a more equal distribution of the national income by imposing heavy taxes on the rich and using the revenue thus obtained to finance public services.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Taxes are compulsory payments, but the prices charged by publicly operated commercial enterprises are not compulsory. Is there any real difference in the matter of compulsion between a tax and the rate charged by a municipally owned water supply?

2. A current newspaper article states that the people of the United States spend more on automobiles each year than they do on all governmental services. It is then argued that we ought to spend more on such services as schools, parks, etc. Is this a good argument? Is it a valid reply to those who assert that the burden of taxation is too great at the present time?

3. As more and more of the total private income of the people of this country is converted into socialized income by means of taxation, which class or classes do you suppose are likely to find that their real incomes are increased: the rich? the poorest class? the lower middle-income class?

4. To take large amounts by taxation from the rich to increase public expenditures is said to be disguised socialism. In which respects, if any, does the heavy taxation of the rich receive support from the doctrines of socialism? In which does it not receive support?

5. A great deal of revenue could be raised from a tax on a vital necessity of life such as salt. Would a tax on salt be a desirable one to introduce into the United States? If not, which of the tests of a good tax does it fail to meet?

6. In some states the motor-vehicle tax varies roughly with the value of the car. Is this "just"? If so, can you justify the tax on the theory of benefit? on the theory of ability to pay?

7. During the World War the United States imposed a tax on business profits in excess of a certain amount. Opponents of this tax asserted that it (1) raised prices, (2) caused extravagant business expenditures for advertising and the like, and (3) discouraged enterprise. Explain how the tax might have caused each of these effects. Show that it could not have caused them all at the same time.

8. "Taxes which are unequal are likely to be shifted. The progressive income tax is an unequal tax. Therefore it is shifted." Criticize.

BOOK V. INTERNATIONAL ECONOMIC RELATIONS

CHAPTER XXXIV

FOREIGN EXCHANGE

In an earlier chapter we explained the determination of the value of money in terms of commodities in general. There we were concerned with the problem of the general price level. In this chapter we shall consider the conditions that regulate the value of money of one country in terms of the moneys of other countries. Concretely, we wish to discover how the value of American money in terms of British or French money is determined. We wish to know this because American money is not usually acceptable in foreign countries. If an importer in New York wishes to buy goods in London or Paris he must know how many dollars the goods will cost. But the sellers will not accept payment in dollars in New York, or at least they would not do so if there were no arrangements by which they could convert these dollars into pounds or francs in London or Paris.

The Meaning of Foreign Exchange. By the term "foreign exchange" we mean the rate at which the money of one country can be exchanged for the money of other countries. Dealers in foreign exchange are engaged in buying and selling the circulating media of various countries. An American traveler about to embark for England wishes to exchange his dollars for pounds sterling, the British monetary unit, for, of course, American money does not circulate in England. How many dollars must he give for a pound sterling? If both countries are on a gold standard, as now, and as they were before 1914, each kind of money represents a certain quantity of gold. The dollar contains 23.22 grains of fine gold, and the pound sterling contains 112.99 grains of fine gold. Since the gold content of the pound is 4.866 times greater than that of the dollar, the American can pay $4.866 for a pound sterling without affecting the quantity of fine gold which he possesses. When two kinds of money exchange at a rate corresponding to their gold content, the rate of exchange is said to be at "par" or at "specie parity."

If the circulating media of all countries consisted exclusively of money, foreign exchange would be merely a matter of exchanging the money of one country for the money of another country.

Forms of Foreign Exchange. In reality the media of exchange in modern countries consist in very large part of "deposit currency." In speaking of deposit currency in previous chapters we have had in mind chiefly the demand deposits of private individuals and business firms. So far as foreign exchange is concerned the deposits of *banks* in foreign banks is of far greater significance, since private individuals rarely have deposits in foreign banks. "Bank drafts" are in reality checks written by a bank against its deposit in another bank. If one goes to a bank to buy "foreign exchange," one will likely get a bank draft or banker's bill, as it is often called, drawn on a foreign bank and calling for the payment of a certain amount of foreign money. Other forms of foreign exchange, such as travelers' checks, letters of credit, and cable transfers, are in reality equivalent to bank drafts, since they are all drawn against deposits maintained by the selling bank or its agent in some foreign bank.

Other forms of credit instruments that figure in foreign trade are commercial bills of exchange. These are drafts drawn by the seller against the buyer for the amount of the purchase and are in no way different from the domestic bills of exchange discussed in previous chapters, with the exception that they are drawn in terms of foreign money. They constitute an order on the part of the maker of the bill to the person or firm upon which it is drawn, to pay the sum of money designated. The buyer of the goods may, however, arrange that the bill shall be drawn not upon himself, but upon a bank with which he has made suitable arrangements. A draft or bill of exchange drawn upon a well-known bank obviously has an advantage over a bill drawn upon a relatively unknown business man or firm.

Foreign-Exchange Practices. As a rule American exporters 1
draw bills of exchange against the foreign buyers (or banks
designated by the buyers). The American exporter takes this
bill of exchange to his bank, has it discounted, and receives a 2
credit on his deposit account. The American bank sends the
bill of exchange to its foreign correspondent bank, which in 3
turn sends it on to the bank or firm upon which it is drawn, for collection if it is a sight draft or for "acceptance" if it is a time draft. If it is a time draft and is presented for acceptance, the person upon whom it is drawn writes the word "accepted" across the face of the bill and signs his name. The bank presenting it for acceptance may then either hold it until it matures, or sell it in the open bill market. Whichever is done, the person or bank holding the bill earns, of course, the difference between the face value of the bill and the discounted value at which it was purchased.

It is particularly important to note the effect of this trans-
action on the deposit account which our American bank carries 4
with its foreign correspondent. When the American bank remitted the bill of exchange to its foreign correspondent, it was credited on the books of the foreign bank (after the bill had been sold in the open market) with the discounted value of the bill. Thus when American banks discount bills of exchange drawn on foreigners and remit them to their foreign correspondents, they are building up their foreign balances. They
are therefore in a favorable position to sell bank drafts drawn 5
against those balances.

Most of the trade between Great Britain and the United
States is settled in terms of British money. As we noted above,
American exporters usually draw their bills in pounds sterling
rather than in dollars. British exporters also usually draw
their bills in terms of British money. The American importer
1 usually instructs the British exporter to draw the bill on some
London bank with which an arrangement has been made. The
American importer must, of course, supply the London bank
with the necessary funds with which to pay the bill when it
falls due. The advantage of having the bill drawn on the
London bank rather than on the American importer is that a
bill so drawn can be sold to better advantage in the open bill
market. When the English exporter has drawn his bill on the
2 designated London bank, he sends it to this bank to have it
3 "accepted." He can then readily sell it in the open bill market
or discount it at his bank, which in turn will either sell it in
the open market or hold it until it matures. A week or so
4 before the bill matures the American importer will go to his
bank and buy a bank draft on London and forward it to the
London bank against which the British exporter was instructed
to draw his bill, thus supplying the bank with the funds needed
to meet payment when the bill matures. In this manner
Americans who have bought goods from British merchants
furnish a demand for bank drafts drawn on London banks.
The American banks will be glad to sell such bank drafts in
order to dispose of their foreign balances.[1]

---

[1] Banks sell foreign exchange even though they carry no balances in London or other foreign centers. They are able to do this by means of arrangements made with the large New York banks which do have foreign balances. A simple illustration will serve to indicate the manner in which this is done. The writer is occasionally obliged to remit a draft to some English bookseller for foreign books or magazines. He usually goes to the University State Bank, Minneapolis, Minnesota, and buys a sterling demand draft. The University State Bank draws the draft directly against the Midland Bank, Ltd., London. The University State Bank carries no balance with the Midland Bank, but it has an arrangement with the Chase National Bank in New York by which it can draw drafts against balances carried by the latter institution in the Midland Bank. The amounts thus drawn by the University State Bank are deducted from its balance with the Chase National Bank. In addition to the amount of the draft the customer is charged a small sum for the service.

Let us sum up the foregoing briefly. The American exporter draws a bill of exchange on the foreign importer or on a designated bank. The American exporter then has this bill discounted at his bank. He is credited on his deposit account with the discounted value of the bill. The American bank then remits the bill to its London (or other foreign) correspondent, and is credited with the discounted value of the bill. Thus American exports serve to increase the foreign balances of American banks. On the other hand, the American importer instructs the English exporter to draw upon some London bank which has by previous arrangement agreed to lend its name. When the bill falls due he goes to his American bank and buys a bank draft to meet payment on the bill. The sale of this bank draft lowers the American banker's foreign balance. Thus American exports increase our foreign balances; imports decrease them. To put it in another way, exports give rise to a supply of foreign exchange, and imports give rise to a demand for foreign exchange.

Not all trade with England is settled in sterling exchange, as has been assumed above. Before the war approximately 90 per cent was so settled. In recent years the percentage is much smaller. There is a considerable amount of the trade between the two countries which is settled in dollar exchange. In this event the British exporter draws on a New York bank, and British exports tend to pile up British balances in New York. On the other hand, the American exporter may be instructed to draw his bill in dollars on some New York bank. When the bill matures, the British importer goes to his bank and buys a draft drawn on New York. Thus, if dollar exchange is utilized, British exports give rise to a supply of bills on New York, and British imports give rise to a demand for bank drafts drawn on New York.

If the exporters of each country followed the practice of drawing their bills in terms of the money of the country in which the goods were sold, the net effect would be that foreign balances would tend to pile up in both countries. Such balances could easily be disposed of by the simple device of the two correspondent banks drawing drafts against each other.

Before the war much of our trade with continental Europe, Asia, and South America was paid for in sterling exchange. London was, and still is, the chief center for foreign transactions.

The Rate of Exchange. *Specie Parity.* Consider now the effect of the trade currents on the rate of foreign exchange. If the foreign purchases made by Americans exactly equal the sales made to foreigners, the American bankers' foreign deposits will always remain constant, and the bankers can afford to continue to buy and sell exchange approximately at par, making their profit by selling at a little higher rate than the one at which they buy. But if Americans sell more to foreigners than they buy from foreigners, the dealers in exchange will be unable to get rid of their foreign balances; that is to say, they will be unable to sell enough bank drafts to keep their foreign balances constant, since every sale by American business men tends to build up these balances and every purchase tends to reduce them. With huge foreign balances the American banks have these alternatives: they may invest these surplus balances by purchasing stocks, bonds, or commercial paper in the foreign country, or they may have the balance transferred to their own vaults by ordering a shipment of gold. The latter obviously involves an expense; and rather than incur this, they would prefer to sell bank drafts below the par of exchange, though not below a point which would cover the expense of shipping the gold. With huge foreign balances exchange rates might then tend to drop to the "gold-importing point," or (in the case of British exchange) to about $4.84. Conversely, if American business men are buying more from foreigners than they sell to foreigners, the demand for bank drafts will be so great that the foreign balances of our banks will prove insufficient. The American bankers will therefore be compelled to ship gold to their correspondent banks, thus building up a balance which will supply the demand. But they will be unwilling to do so unless they can sell drafts at a price that will defray the cost of shipping. Thus exchange will rise to the "gold-exporting point," or to about $4.89. Competition among dealers will normally prevent it from going higher.

When both countries are on a gold basis the rate of exchange will then be somewhere between the "gold-shipping points." Exchange rates do not, of course, jump at once from par to one of the gold-shipping points. If the trade currents indicate that the foreign balances will prove excessive, the rate of exchange will gradually drop below par; if, on the other hand, they indicate a probable deficiency in the foreign balances, the rates of exchange will gradually rise.

American purchases of merchandise or services from foreigners tend, then, to raise the rate of exchange, whereas sales to foreigners tend to lower it. It must be remembered that we buy from abroad not merely merchandise imports but also shipping services, banking services, the pleasures of travel, securities, and the services of invested capital.

If one country is on a silver basis and another is on a gold basis, the par of exchange will fluctuate according to the changes in the market rates of gold and silver. In other respects the forces determining the rates of exchange are similar to those discussed above.

Purchasing-Power Parity. Let us now consider the exchange rate between two countries one of which is on a paper basis, as was true in the United States from 1862 to 1879 and as was true of most European countries during and immediately after the World War. In such a situation the exchange rates may depart very far from the specie parity. The normal rate of exchange will then be the ratio of the purchasing power of money in the two countries. If, for example, the British paper pound were able to buy twice as many goods as the American dollar, the normal rate of sterling exchange would be $2. This rate would in this particular case be the purchasing-power parity.

Let us suppose that the British exchange rate is temporarily far below the purchasing-power parity. In this situation Americans can exchange dollars for pounds sterling, which will buy much more than the dollars so exchanged. This will stimulate imports into the United States from England. The movements of goods from England into the United States will tend to raise prices there and lower them here, and this process will continue until the price levels of the two countries correspond to the exchange ratio.[2] On the other hand, the increase in imports will result in an increase in the demand for bank drafts on London, with a consequent tendency on the part of exchange rates to rise. In this way, if the exchange rates are out of line with the relative purchasing power of the two kinds of money, trade currents are set in motion which tend to adjust the exchange rate to the two price levels.

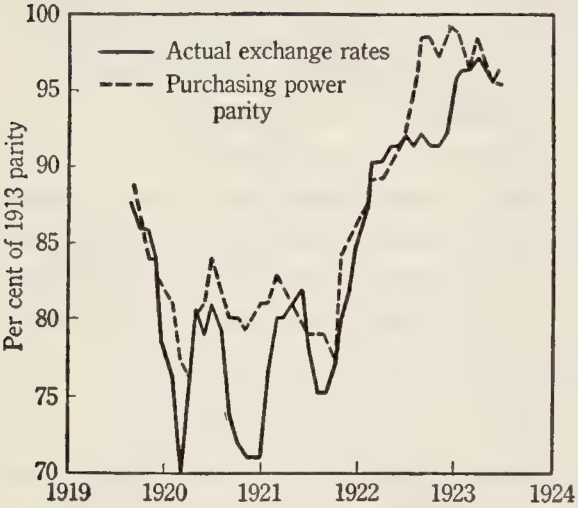

FIG. 49. Exchange rates: United States; Great Britain[1]

It is, in part, because of the interaction just noted (which operates also when both countries are on a gold basis) that the payment of reparations or war loans must in the long run result in larger exports from the countries making these payments. Let us take the case of the reparation payments. To make these payments the German government must raise the revenue, in part, through taxation. With the funds thus secured the Agent General for Reparation Payments then goes into the market to buy foreign bills of exchange, which will be sent on to the Allied governments. The result is that Allied exchange is bid up; for example, relatively more marks will be offered for a given number of francs. From the French point of view this means a decline in exchange rates. The exchange rate is forced out of line with the purchasing-power ratio of marks and francs. Francs converted into marks will buy more goods than the francs will buy in France. Germany thus becomes a desirable place in which to buy, and exports from Germany are stimulated.

---

[1] See J. M. Keynes's *Monetary Reform*, p. 112.

[2] This is based on the assumption that the prices of goods which enter into international trade fluctuate approximately in accordance with the general wholesale-price level. This is not true, strictly speaking, but on the whole there is a tendency for them to move together, since general price changes due to monetary causes affect all commodities more or less equally. In periods of great price upheaval, however, there are almost always certain shifts in demand and production which upset to some extent the usual price relationship. Cf. Gustav Cassel, *Money and Foreign Exchange after 1914*.

Thus the reparations payments tend to make the exchange rates favorable to the Allied countries. Moreover, these payments will also affect the relative price levels of Germany and the Allied countries. The Germans will be taxed by their government for many years to pay the reparations. This will reduce the income of the German people.[1] They will be unable to buy as many goods as they would if they were not taxed so heavily. As a result of this reduction in their money incomes and the consequent decline in their purchasing power, the prices of commodities will fall in Germany to a level below that in other countries, exchange rates considered. From the standpoint of other countries Germany will therefore be a good place in which to buy. The Allied countries, moreover, will have a larger purchasing power because of the reparations payments. Prices will therefore tend to rise in these countries. In this situation the Allied countries will constitute an excellent market for German goods. As goods flow out of Germany the *decline* of prices there will be checked, and as goods flow into the Allied countries the *rise* in prices there will be checked. Thus prices tend again to be brought back into line with the exchange rates. A stream of exports sufficient to maintain this equilibrium will continue so long as reparations payments are continued unless these payments are balanced by new debts.

---

[1] The Agent General for Reparation Payments uses the credits secured from taxation to buy the foreign drafts. The government deposits are therefore reduced, the effective volume of circulating media is curtailed, and so prices fall. Taxation for domestic expenditures would not have this effect because the funds thus raised would be used to buy goods within the country.

The payment of the reparations therefore tends to stimulate German exports for two reasons: first, such payments tend to make German exchange rates favorable to the Allied countries; second, they tend to lower German prices. So long as prices are out of line with the exchange rates, exports will tend to flow out of Germany. Equilibrium of prices and exchange rates can only be reëstablished by a volume of German exports (in excess of normal) equivalent to the reparations payments. It should be noted, however, that these are the long-run tendencies. Owing to the large sale of German securities abroad there has not yet developed an excess of merchandise exports in spite of the reparations payments. This situation will be explained in some detail in the following chapter.

Speculative Rates. To return to the main argument, we have still one more possible situation to consider. Let us take the exchange situation between two countries one of which is so situated that not only is it impossible to obtain gold and ship it out but it is also impossible to ship goods out of the country to any considerable extent. During the war and for a considerable period after the war it was practically impossible, for example, to get any goods out of Germany. Suppose an American banker had a balance in Germany either carried over from a previous period, when trade was open, or obtained from the sale of securities or other property owned in Germany. Under the conditions named what could he do with his balance? He could obviously sell it to someone else, but what could that person, in turn, do with it? He could not demand gold and ship it to this country; he could not go out and buy goods to ship to this country and thus indirectly, through the purchase, shipment, and sale of these goods in the United States, convert his marks into dollars. If neither gold nor goods could be got out of Germany and shipped to this country, there would be no possible way of getting this balance out of the country. Only one possible thing could be done with the balance: it could be spent or invested in Germany. But at such a time very few foreigners would be willing either to go there and spend it or to risk an investment in German property with all sorts of uncertainties coming up, such as the danger of revolution, the destruction of property values, confiscatory taxation, and continued issues of paper money. Clearly any person who bought those German marks would be gambling with the future. Yet there are always people who are willing to take great risks in the hope of great gains. So there would probably be a sale for the marks; but the price would have nothing tangible, such as gold or goods to tie to. The exchange rate would be a purely speculative one, based on the personal judgments of the buyers and sellers as to the future of the country in question.

We may, then, have three possible situations. First, the situation in which there is free movement of specie between the two countries. The normal exchange rate in this case corresponds to the ratio of the metal content of the two kinds of money. Second, we have the situation in which there is free movement of goods but not of specie. In this case the normal exchange rate corresponds to the ratio of the purchasing power (in terms of commodities entering into international trade) of the two kinds of money. Third, we have the situation in which there is free movement of neither gold nor goods. In this case there is no normal for exchange rates; since there are no trade currents which can be set in motion to establish any normal level, the exchange rates will be on a purely speculative basis.

In the first two cases mentioned the exchange rates often depart temporarily from normal, but the movements of gold or goods which follow tend to eliminate the divergence and bring the exchange rates back to normal.

Pegging the Rate of Exchange. When the exchange rates are unfavorable, importation of foreign goods becomes very difficult, but national interests may require that imports be maintained. A nation may then decide to regulate artificially the rate of exchange. During the last years of the World War, Great Britain decided to "peg" the rates of exchange in New York. This means that a point was established ($4.70) below which the rate was not allowed to fall. This could only be done by artificially creating a sufficient demand for bills to absorb the entire supply offered at this price. Great Britain undertook to purchase at $4.70 all the bills offered, thus making it impossible for the rate to fall below this figure. The British government obtained the funds needed to make these purchases from loans granted to it by the United States Treasury.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Is there any good reason why a formal organized market such as the stock exchange has not been established in which foreign exchange is bought and sold?

2. Do finance bills (that is, bills drawn by banks in excess of their foreign balances, and, therefore, constituting in effect a method of borrowing from abroad) have any influence in steadying the rate of exchange? Explain.

3. Some bills of exchange are of such unquestioned reputation that no security in the nature of accompanying documents are necessary. Other bills, however, known as *documentary bills* have certain documents attached, such as a bill of lading, an invoice and a certificate of insurance. These documents, without which the importer cannot obtain possession of the goods, will not be released until the bill has either been paid or "accepted." In what respect, then, are documentary bills superior to clean bills?

4. Indicate how each of the following would tend to affect the rate of exchange in New York on London: exports of merchandise; American tourists traveling abroad; sale of American securities abroad; payment by the British government of a part of its debt to the United States; payment of interest on foreign bonds owned by Americans; payment by America to foreign shipping companies.

CHAPTER XXXV

THE BALANCE OF TRADE AND INTERNATIONAL DEBTS

It is a common error to suppose that it is especially desirable for a country to export more than it imports. This is an ancient fallacy. The mercantilists considered an excess of exports to be a "favorable balance of trade." We still use the term "favorable balance of trade" to indicate an excess of merchandise exports and "unfavorable balance of trade" to indicate an excess of merchandise imports, but we no longer believe that the terms have any significance. We do not admit that either one is necessarily more favorable than the other; rather, every country tends to pass through several stages in its development, and in certain periods a favorable balance of trade, so called, is the normal thing, whereas in others an unfavorable balance will naturally obtain.

Our Balance of Trade before 1873. Up to 1873 the United States had an "unfavorable" balance of trade. This followed from the fact that up to this period we were a borrowing country. We had abundant natural resources and numerous opportunities for profit-making, but little accumulated capital. All industries were greatly in need of capital; consequently we were constantly selling industrial securities abroad. The state and national governments also borrowed heavily in this period, the latter particularly during and following the Civil War. The effect of floating these loans in European countries was to lower the exchange rate below the normal or purchasing-power parity, since the sale of securities abroad would tend to increase American bankers' balances abroad and thus give rise to a larger supply of sterling exchange. With the exchange ratio below the purchasing-power ratio it was advantageous for Americans to convert their dollars into pounds sterling and to buy exportable merchandise in England and ship it into the United States. The sequence of events therefore was as follows: the sale of securities abroad depressed the exchange rate below the purchasing-power parity. Low exchange rates stimulated imports into the United States. On the other hand, low exchange rates discouraged exports; for while the price received for American exports was established in the foreign markets in terms of foreign money, the American exporter received fewer dollars because of the fall in exchange. The increase in imports and decrease in exports tended to rectify the exchanges, but with fresh borrowings the exchange rate again tended to be depressed. As a matter of fact, particularly in the decade prior to 1873, foreign exchange was almost constantly depressed below the purchasing-power norm.

A country which is borrowing heavily is therefore almost certain to have an excess of imports. The excess of imports follows from the fact that it is a borrowing country. If it is desirable for such a country to borrow capital, it can scarcely be said that the resulting excess of imports can properly be referred to as an "unfavorable" balance of trade. Yet we technically so refer to it because that has become for historical reasons the settled usage.

In one respect, however, the increase in imports and the decrease in exports has an unfavorable transitory effect on industry. It results in a shifting of capital and labor from the exporting industries to industries producing for the home market. In the long run it is, of course, advantageous to have labor and capital flow from the less advantageous industries to the more advantageous until the relative profitableness of different industries has been equalized, but the transition involves an economic loss for both labor and capital. Thus while the increase in imports and the decrease in exports could not be said to be disadvantageous to the country as a whole, it is certainly true that such a shift affects unfavorably that portion of the labor and capital which is engaged in the exporting industries. But this holds true only for the transition period; in the end the redistribution of labor and capital will equalize the relative profitableness of exporting as compared with domestic industries.

Our Balance of Trade as Debtor Country from 1873 to 1914. In 1873 the volume of loans placed by the United States in Europe reached a peak. Subsequent to that year the borrowing abroad rapidly declined, and by 1876 the repayment of these loans began. Moreover, there was the interest to be paid on the past borrowings. From 1874 to 1914 the United States could therefore properly be called a debtor country, a country which had already borrowed extensively abroad and was now in process of paying interest on past borrowings and repaying past loans. Of course it must not be inferred that in this period Americans placed no fresh loans in Europe, but the payments on past loans overshadowed the current borrowings.

The payment of interest on past loans as well as the repayment of loans tended to raise the rate of foreign exchange, since the debtors — persons or corporations as well as the United States government — were compelled to go into the market and purchase bank drafts drawn on foreign centers with which to pay these obligations. The demand for foreign exchange thus increased and the rate of exchange tended to rise. With the rate of exchange above parity, exporting became more profitable, since at the same foreign prices American exporters got a larger number of dollars for merchandise sales abroad than for sales at home. On the other hand, importing became less profitable, since it now required more dollars to buy the same foreign goods because of the higher exchange rate. Therefore, with the cessation of borrowing and the consequent rise in exchange rates, our balance of trade reversed itself as compared to the period prior to 1874. Merchandise exports now exceeded merchandise imports. We had a so-called favorable balance of trade.

Labor and capital now tended to be diverted toward the exporting industries, since these were relatively more profitable than the industries producing for the home market. In a large country like the United States, with a relatively unimportant foreign trade, this probably did not result in any actual movement of capital and labor from the domestic industries to the exporting industries. A relatively larger proportion of new capital and the growing population probably went into the exporting industries than would otherwise have been the case. Thus from the standpoint of the exporting industries the transition period from an unfavorable balance of trade to a favorable balance is genuinely advantageous, though from the standpoint of labor and capital as a whole no such advantage is evident. And in the long run the temporary advantage enjoyed by the exporting industries soon disappears as more labor and capital are diverted into these industries. The relative profitableness of the domestic and exporting industries is thus rapidly equalized.

Increased Exports during the World War. With the coming of the World War the United States passed into a third stage. Prices rose rapidly in European countries compared with the American price level. Europeans therefore found it to their advantage to purchase huge quantities of American exportable products. In other words, since prices in Europe had risen more than had dollar exchange, Europeans could buy more cheaply in America than in Europe. American exports increased enormously. Export industries were enjoying an immense prosperity, and more labor and capital were diverted into the production of those goods which were demanded by the European nations. Not only was the United States a good place in which to buy goods, but at the same time it was an excellent place in which to market securities. Interest rates rose rapidly in European countries and ruled higher than in the United States. Large blocks of foreign and American securities were resold in the United States, and foreign securities as well were disposed of to American investors. Our excess of exports was balanced in part by this huge purchase of foreign securities.

The United States as a Creditor Nation. Before the war the total American securities held abroad amounted to about five or six billion dollars; during the war the greater part of these securities were purchased by American investors, thus practically wiping out our former indebtedness. Moreover, during the war the United States government lent some ten billion dollars to the Allied governments. In addition private investors have during the war and post-war years purchased some twelve billion dollars of foreign securities. Hence we are now a creditor nation to the extent of about twenty-two billion dollars, second only to Great Britain, which is still a net creditor nation to the extent of about twenty-five billion dollars. In the very near future we shall be the largest creditor nation in the world. Our annual purchases of foreign securities are still growing and are reaching large proportions. From 1921 to 1923 about half a billion dollars of foreign securities were placed in the United States annually; from 1924 to 1926 a billion dollars were sold each year; and in 1927 over a billion and a half were disposed of in this country. But this was offset by the repayment by foreigners of past loans and the purchase of new securities. Thus in 1926 our net investment abroad was only about six hundred million dollars, and in 1927 about six hundred and seventy-five million.

It is largely because of these purchases of foreign securities by American investors that we continue to maintain an excess of merchandise exports over imports. Before the war this excess amounted to approximately half a billion a year; in the post-war period from 1921 to 1924 it averaged about a billion dollars a year; from 1925 to 1927, however, the excess has declined to an average of about half a billion a year.

Table XXXVIII gives our balance of trade for the year 1927. In this year the net excess of merchandise and silver exports was about $500,000,000. This was more than balanced, however, by our tourist expenditures abroad. The balance of payments in our favor arising from interest payments amounted to about $500,000,000, but this was more than offset by the net balance against us of securities purchased.

Table XXXVIII. Balance of Payments of the United States for 1927, in Millions [1]

Exports of Merchandise and "Invisible Exports"		Imports of Merchandise and "Invisible Imports"	
Exports of merchandise and silver	$5037	Imports of merchandise and silver	$4489
Tourists	153	Tourists	770
Freight and passenger traffic	229	Freight	172
Interest receivable	795	Interest payable	281
War-debt receipts	206	Immigrants' remittances	241
Miscellaneous	291	Miscellaneous	259
New Foreign Investments in the United States	919	New American Investments abroad	1648
"Backwash" on previous American investments abroad	767	"Backwash" on previous investments here	709
Net Exports of Gold	166		
	$8563		$8569

Looked at in a broad way, the main facts are as follows: our purchases of goods and tourist pleasures approximately balance our sales of goods to foreigners. For the rest the heavy interest charges which foreigners owe us each year are balanced by further borrowing. The interest on old debts is paid by contracting new debts.

Eventually, however, the interest payments on past debts will become so great that they must exceed the new debts currently contracted. When that time comes the United States will have passed into the stage of a mature creditor nation. The interest on her old loans will overshadow the new loans. This is the situation in which England has found herself since 1853. When the United States reaches that stage her imports of goods and services will inevitably exceed her exports. Confronted with these heavy interest charges, Europeans and other foreigners will be forced to go into the market to buy dollar exchange. But the payments due them from abroad will not equal the payments they must make abroad. The demand for dollar exchange will therefore exceed the supply at the par rate of exchange. The exchange rate will therefore become unfavorable to Europeans and favorable to the United States. This, of course, means that the exchange rate will be out of line with the purchasing power of the American dollar. As a result, American dollars will buy more goods in Europe than they can command in the United States. Europe becomes a good place in which to buy. Imports into the United States will increase. On the other hand, with an exchange rate against them, Europeans cannot so profitably buy goods in the United States. American exports will therefore tend to decline. As a result, merchandise imports will tend to exceed merchandise exports. We shall inevitably have an unfavorable balance of trade.

---

[1] *Trade Information Bulletin*, No. 552, United States Department of Commerce.

Economic Effects of an Unfavorable Balance, So Called. But will this excess of imports really be unfavorable, properly speaking? It must be obvious from the foregoing discussion that the popular alarm on this point is not well founded. To be sure, export industries may suffer for a time; but it will not be long before labor and capital will have been diverted sufficiently to equalize the relative profitableness of export as compared with domestic industries.[1] Specifically, this probably means a relative decline in agriculture and other industries producing raw materials for export, and a relative increase in manufacturing.

At this point a further doubt is almost certain to arise. Will not the excess of imports injure the manufacturers of the United States? Those industries producing for the home market will not merely have to face the competition of the labor and capital diverted from the foreign market because of the relative unprofitableness of the exporting industries, but they will also have to face the competition of the large imports from foreign countries. Will not the Americans lose their domestic markets as well as their foreign markets?

---

[1] Labor and capital will also tend to be diverted somewhat from those domestic industries which will be forced to compete with foreign imports to industries not subject to foreign competition. The shifts in production will, moreover, be affected by the direction of the new demands created by the enlarged purchasing power derived from our position as a creditor nation.

It is clear that the increase in imports and the diversion of labor and capital from the export field will tend to lower the prices of goods sold in the United States. What is feared is that prices will be forced so low in this country that American entrepreneurs will be forced out of business; but even a brief examination of the facts will show that this view is entirely untenable. If our imports exceed our exports, as they necessarily must in the long run if we continue to be a creditor nation, it must necessarily follow that this country will be a favorable place for foreigners to sell in or they will certainly cease to export to this country. It must also be true that this country will continue to be a favorable place for our own producers to sell in or they will certainly divert their production to foreign markets. In short, our price level must remain relatively high compared with European price levels, foreign exchange rates considered. In other words, the rates of exchange as compared with the purchasing-power ratios must be such that the American market is sufficiently superior to the foreign markets to bring in the excess of imports over exports which our position as a creditor nation requires. The fear that American prices under the competition of these imports will drop in relation to European prices (exchange rates considered) is therefore entirely unfounded. In fact, the imports would not continue to come in were it not that the American market was slightly superior, prices and exchange rates considered; and this superiority is due to the increased purchasing power which a creditor nation necessarily has.

N.B.

To put the matter in a somewhat different way: it is feared that the excess of imports will displace American labor and capital and prevent their productive employment. But if we include in the terms "imports" and "exports" not merely merchandise but also services, omitting only loans and the interest and dividends payable thereon, it follows that in the *long run* the excess of imports can only equal the payment made on loans either in the way of repayment of principal or in the payment of interest and dividends. If the imports exceeded this amount the demand for foreign exchange would cause exchange rates to rise above par and as a result imports would be checked and exports would be encouraged; in other words, the excess of imports over exports can never in the *long run* exceed the additional purchasing power which Americans become possessed of because of the repayment of loans or the payment of interest and dividends. The remaining imports are paid for by the purchasing power derived from our exports, leaving as much purchasing power as before with which to purchase American-made goods.

There is, therefore, never any danger from a long-run point of view that foreigners will be able to destroy our home market by excessive imports. Nor is it true that we cannot keep our labor and capital employed unless we can dispose of our "surplus products" in foreign markets. "Surplus products," as popularly used in this connection, apparently means products which cannot be sold at home. There are, in fact, no such surplus products. No nation ever produced goods that it did not in the end purchase itself directly or indirectly. If a country exports more than it imports, it is always either in the process of increasing its investments abroad or it is a debtor country whose payments on interest exceed its new borrowings. If it is increasing its investments, the exports will exceed the imports by the amount the new loans exceed the interest on the old loans. It is not difficult to see that such a country itself supplies the purchasing power which buys these so-called surplus products. The case is merely that the people of such a country prefer to invest part of their income rather than spend it, and the comparative rates of interest are such (exchange rates considered) that it is more profitable to invest in foreign countries than in their own country. Such an exporting country purchases its own "surplus." If the country is a debtor country, it cannot properly be said that it really has produced the exported "surplus." Rather the capital owned by foreigners in this country produced it, and, in effect, it is sold abroad to meet the payment of interest and dividends on this foreign capital. Conversely, if a country has an excess of imports it is either a borrowing country or a creditor country. If it is a borrowing country, foreigners are really furnishing the purchasing power with which the excess is bought. Properly speaking, the foreigners are the real purchasers of this excess of imports. If the country is a creditor country, its total production is to be measured not merely by what is produced on its own soil but also by the product attributable to the capital invested in foreign countries. Every nation's purchasing power is equal to its total production. Every nation ultimately buys its own product. In reality we never sell any surplus to foreigners. If we consume less than we produce, we ourselves buy that surplus, which may be invested either at home or abroad; if we consume more than we produce, we are living off our capital. We may, however, consume more than we produce on our own shores, the excess being produced by our capital invested abroad. That is the situation in which creditor nations generally find themselves. Creditor nations are rich nations; and from that point of view it may be said that creditor nations are the most favorably situated, yet they regularly have (unless they are making new loans which more than counterbalance the effect of their past loans) an "unfavorable" balance of trade.

Factors governing the Balance of Trade. Nor is the fear well founded that purchases of foreign goods, services, or securities "will drive money out of the country." If such purchases exceed our sales to foreigners, the demand for foreign exchange may, it is true, rise above the gold-shipping point (both countries are assumed to be on a gold basis). Certain forces are set in motion, however, which check the flow of gold out of the country. In the first place, the high rate of exchange will tend to stimulate exports and discourage imports and thus put a stop to the excess of purchases over sales which originally produced the tendency. In the second place, the outward flow of gold will in time reduce the bank reserves, raise temporarily the rate of interest, and make this country a favorable one in which to invest capital (that is, to buy securities); or, to put it another way, the high rate of interest will lower the prices of securities in this country, making it a favorable country in which to buy stocks and bonds. But such sales of securities to foreigners will result in the drawing of a large volume of bills on foreigners, will increase the supply of foreign exchange, will lower the rate of exchange, and will stop the flow of gold. In the third place, the outward flow of gold will in time cause a contraction of the circulating media of the country, which in turn will result in a fall in prices, making this country a desirable place in which to buy. The consequent increase in exports will give rise to an increased supply of foreign bills, American bankers' balances abroad will be increased, rates of exchange will fall, and the flow of gold will be checked.

Thus the factors that regulate the currents of foreign trade are (1) the rate of exchange, (2) the price levels at home and abroad, and (3) the rates of interest in the respective countries. If the price level is higher in one country than in another (by which we mean that the purchasing power of its money, in terms of goods, is less than the purchasing power of the foreign money, exchange rates considered), goods will tend to flow into the country where the price level is high. But this movement of imports will itself tend to rectify the divergence between the rate of exchange and the price level. In like manner, if the rate of interest is relatively high in one country, investments will be made in that country and securities will be purchased. Such purchases will tend to rectify the divergence between the exchange rates and the normal relation of interest rates in the two countries.

The Effect of Reparations and War Loans on International Trade. A word may here be added with reference to the effect of the payment of reparations or huge war loans on the currents of international trade. Take the case of the debt that Great Britain owes the United States. When Great Britain makes payments on this loan, she goes into the market and buys dollar exchange to send to our Treasury Department at Washington. The effect of these huge purchases of dollar exchange is to raise dollar exchange in London and to lower sterling exchange in New York. The exchange rate therefore becomes unfavorable to Great Britain and favorable to the United States. In this situation it will be profitable for Americans to buy exportable goods in Great Britain, ship them to the United States, and pay for them by buying exchange at the favorable rate; in other words, the dollar will buy more in Great Britain than in the United States. The effect of the repayment of the war loan will therefore be to stimulate imports into the United States and discourage exports. Will this injure American industries? As has been explained above, it will obviously injure the export trade as well as those industries which encounter British competition in our domestic trade unless Americans invest an equivalent amount in foreign securities (not necessarily British), making payment for these in pounds. If such investments were made, the exchange rate between England and the United States would not be affected, nor would our balance of trade be affected. It is not necessary, however, that we should balance the repaid loans with equivalent foreign investments; we could increase our total consumption as a result of the increased purchasing power derived from the repaid loans. With this increased purchasing power we could buy the excess of imports from Great Britain. The situation is analogous to that of a creditor country. While industry as a whole would not in the long run suffer from this situation, it is clear that the drop in the exchange would unfavorably affect those American industries which are subjected to British competition, and if the exchange rate were continuously depressed for many years by the prolonged repayment of these loans, these industries would be less profitable than those not subject to British competition;[1] in fact, they might become so unprofitable that labor and capital in these industries would have to be diverted to other uses. We may therefore expect that such industries would cry out against the repayment of the debt as an economic loss; and since the adversity of certain industries is more conspicuous than the general prosperity of industry as a whole, they might succeed in making people believe that the country was suffering from the repayment of the loans. Any country which receives large payments of reparations or war loans places itself in the position of a creditor nation and must expect an excess of imports unless it is at the same time making correspondingly large foreign investments. Certain industries are adversely affected by this flood of imports. The necessary readjustment of labor and capital is, in fact, an injury. The situation is analogous to the tariff situation. Free trade may be economically superior to a protective system; but if we suddenly take away the tariff, the labor and capital now employed in the protected industries will have to go into other lines of production. The transition period is a difficult and costly one. In like manner, the payment of reparations or war loans may result in a redistribution of capital and labor which, from the immediate point of view, is an evil.

In so far as the payment of war reparations or war loans seriously upsets the exchange rate, the immediate effect is similar to that of a heavily declining price level. This may be illustrated by reference to the situation in England in the autumn of 1921.[2] Germany was called upon to pay one billion gold marks of reparations in cash. The effect of this payment was to depreciate the mark. As a result, German exporters were able to quote extraordinarily low prices in England. British producers became alarmed and reduced output. Trade was paralyzed, not because German imports actually entered in any large volume but because of the uncertainty of prices caused by the depreciation of the German mark. If the reparation payments could be made so gradually that the exchange rates would not be thrown violently and suddenly out of line with the purchasing-power parity, no such disastrous effect would follow; if the exchange rates were thrown but slightly out of line, the currents of international trade would gradually be adjusted to the situation in the manner that we have described. Just as a severe decline in the price level throws industry into depression, so a sharp depreciation of exchange rates caused by heavy reparation payments paralyzes trade. But, on the other hand, just as a slowly receding price level permits the gradual readjustment of wages, returns on capital, rents, salaries, etc., and thus does not seriously interfere with industrial prosperity, so a moderate divergence of exchange rates from the purchasing-power parity will make possible a gradual adjustment of the currents of international trade to meet the new situation without serious interference with established industries. The great difficulty with war indemnities, reparations, and loans arises when they are very large and have to be paid within a relatively short period of time. Then they can scarcely fail to disturb seriously the established equilibrium in international trade. It is in this respect that they differ sharply from the normal payments of interest and principal growing out of the normal currents of international trade.

---

[1] Certain of these industries might, however, be fortunate enough to produce commodities which would be in greater demand in the domestic market because of the enlarged purchasing power consequent upon the nation's becoming a creditor.

[2] See W. T. Layton, "European Recovery," *Proceedings of the Academy of Political Science,* January, 1925.

War-Debt Settlements.[1] As a result of loans made by the United States government during the war and after the armistice, Great Britain, France, Italy, Belgium, and a number of smaller countries are indebted to the American government to the amount of approximately $11,000,000,000.

The principal sum owing to us from Great Britain is $4,600,000,000. According to the agreement reached with that country on June 19, 1923, annual payments are to be made amounting to about $160,000,000 a year for ten years, and thereafter approximately $180,000,000 a year for an additional period of fifty-two years. Belgium owes the United States about $418,000,000. Under the agreement of August 18, 1925, she agrees to begin with an annual payment of $5,000,000 a year, which will increase to about $13,000,000 in ten years. Italy owes the United States the sum of $2,400,000,000 and agrees, under the settlement of November 15, 1925, to pay $5,000,000 a year at the start. The annual payments will increase until by 1936 they will amount to $23,000,000 a year and eventually to $50,000,000 a year. The French debt to the United States is approximately $4,200,000,000. The proposed settlement, which, however, has not yet been ratified, provides for annual payments beginning with $30,000,000 a year and increasing until by 1943 they will amount to $125,000,000.

---

[1] See "The Interallied Debts" (*International Conciliation*, May, 1927), especially the articles by F. W. Taussig and J. W. Angell.

The annual payments for all countries run through a period of sixty-two years. All told, the annual payments average about $210,000,000 for the first five years, $250,000,000 for the next five years, $350,000,000 for the next forty-five years, and $400,000,000 for the last decade of debt payments. At the end of sixty-two years, if these schedules are carried out, some $22,000,000,000 will have been paid. The total payments will therefore be more than twice as great as the principal sums, the remaining sums being interest. But the rate of interest to be paid by the different countries varies greatly. The average rate over the whole period of sixty-two years is 3.3 per cent for Great Britain. For the other countries it is much lower: 0.4 per cent for Italy and 1.6 per cent for France. The great difference between the various settlements can be seen from the following data. If we assume that 3 per cent is a fair average interest rate over the sixty-two-year period, Great Britain has been overcharged 7 per cent, while 28 per cent of the Belgian debt has been canceled, 32 per cent of the French debt, and 62 per cent of the Italian debt; if $4\frac{1}{4}$ per cent is considered a fair average rate, 17 per cent of the British debt has been canceled, 46 per cent of the

Belgian debt, 52 per cent of the French debt, and 75 per cent of the Italian debt.

The total annual payments during the first ten years will average only $233,000,000. This amounts to but one fourth of 1 per cent of our annual national income and only 6 per cent of our annual Federal revenue. It is therefore evident that although the sums are large in toto they are not of great significance when viewed from the standpoint of the annual payments. It is true that the annual payments will reach $350,000,000 after the first ten years; but, even so, the annual payments are relatively insignificant from the standpoint of the United States. But on the other hand, they amount to a very considerable burden on the impoverished European countries.

The Dawes Plan and the German Reparation Payments. After years of discussion with respect to the amount of reparation which Germany should be required to pay to the Allied countries in consequence of the war and the manner in which these payments should be made, the Dawes plan was finally agreed upon in 1924. Under this plan Germany is to make annual payments to the Allies amounting in the first year to $250,000,000 and increasing until in the fifth year the standard payment of $625,000,000, or 2,500,000,000 gold marks, is reached. After 1930 there must be paid, in addition to this standard annual payment, a supplementary sum depending upon the increase in Germany's prosperity. To determine the amount of the supplementary sum an index of prosperity was devised consisting of (1) exports and imports, (2) budget receipts and expenditures, (3) railroad traffic, (4) consumption of sugar, tobacco, beer, and alcohol, (5) consumption of coal, and (6) growth in population.

The plan provides that the funds from which the annual payments shall be made shall be derived (1) from the government revenue, (2) from railway bonds, and (3) from industrial bonds.

The last two sources need some explanation. The German railways have a capital value of 26,000,000,000 gold marks. Before the war these railways were encumbered with debts, but the debts were extinguished by the depreciation of the mark. Toward the end of 1923 it required 10,000,000,000 marks to equal one gold mark. Since these debts ran in terms of marks and not in terms of gold marks, it follows that a prewar debt of 10,000,000,000 marks was now reduced to one single gold mark. The railroads were therefore relieved of their previous indebtedness. The Dawes plan now resaddled the German railways with a bonded indebtedness of 11,000,000,000 gold marks bearing 5 per cent interest and 1 per cent sinking fund. Interest and sinking-fund charges on these bonds would therefore amount to 660,000,000 gold marks a year.

German industrial concerns had similarly discharged their bonded indebtedness by nominal payment in depreciated currency. The Dawes plan accordingly resaddled these industries with a bonded indebtedness amounting in all to 5,000,000,000 gold marks. This is less than the total debt of the German industrial concerns before the war. These bonds similarly pay 5 per cent interest plus 1 per cent for sinking fund. Therefore the annual receipts from these bonds amount to 300,000,000 gold marks. Thus the combined annual payments from the railroad and industrial bonds will amount to 960,000,000 gold marks. The remaining 1,540,000,000 will be obtained from a transport tax and from the general revenue of the government.

Under the plan the 16,000,000,000 railroad and industrial gold-mark bonds become the property of the Reparations Commission, which may, however, sell them in the financial markets.

The Dawes plan separated sharply two distinct though related questions — the *payment* of reparations and the *receipt* of reparations. The first is an internal problem, a question chiefly of adequate taxation by the German government so that the annual payments can be made to the Reparations Commission. The second problem is international in scope. It involves the transfer of the funds to the Allied governments, and this transfer affects international trade and exchange rates. It is one thing to tax the German people and pay the Reparations Commission in German marks; it is quite a different thing to convert these marks into foreign currencies.

The payments are accordingly made in gold marks and deposited to the credit of the Agent General for Reparation Payments. These deposits must then be transferred to the Allies by the purchase of foreign exchange. But heavy purchases of foreign exchange are likely to upset the exchange rates and throw them out of line with purchasing-power parities. If the payments by Germany cannot be transferred without disrupting the exchange rates, the Dawes plan provides that the payments may be allowed to accumulate in the new Bank of Issue, which, under the Dawes plan, is given the exclusive right to issue bank notes in Germany. However, this accumulation of funds in the bank may not exceed 2,000,000,000 gold marks. Surplus funds above this 2,000,000,000-mark limit must be invested in bonds or loans in Germany. But not more than 3,000,000,000 marks may be so invested. If the accumulated payments exceed this total of 5,000,000,000 marks, then the annual payments are to be reduced "until such time as the transfers to the Allies can be increased and the accumulations reduced below the limit named."[1]

In point of fact the funds cannot possibly be converted into the currencies of the Allied countries without upsetting the exchange rates, unless Germany's "earnings from abroad" are "equal to the payments she must make abroad, including not only payments for goods imported but sums paid in reparations";[2] that is, unless Germany's exports of goods and services exceed her imports by the amount of these annual payments. To be sure she may for a time balance these payments by fresh borrowings, but that amounts to paying one debt by contracting a new one. Such borrowings have, in fact, taken place on a large scale. There has thus been an ample supply of foreign exchange with which to effect the transfer of payments. But except for the deliveries in kind, there has been no *real* transfer of payments; that is, there has been no surplus of German exports. On the contrary, Germany's foreign borrowings have exceeded the annual reparation payments to such an extent that her imports have exceeded her exports. These loans have been made not only to various governmental units in Germany, but also to her industries. The prevailing high interest rate in Germany (about 8 per cent) indicates that her industry requires a vast amount of new capital. This flow of capital into Germany is likely to continue until the abnormally high interest rate in that country falls to a more normal level. As that time is approached, Germany's borrowings abroad will gradually decline and eventually disappear. Her annual net importation of capital will thus gradually decline until it is equal to and finally less than the annual reparation payments. Thus gradually a surplus of exports will become possible.

---

[1] Dawes Report (*World Peace Foundation,* Vol. VI, No. 5, p. 379).

[2] Ibid. p. 367.

When this time arrives, however, if the Agent General for Reparation Payments waits (as seems to be contemplated by the Dawes plan) to transfer the payments until the German exports have already reached a point sufficient to allow such transfer without upsetting the exchange rates, it is doubtful if the payments can be transferred at all. If the Agent General hesitates to transfer the payments, allied purchasing power will not be able to absorb a surplus of German exports. A bold transfer of the payments would, in point of fact, only temporarily disturb the price levels and the exchange rates. For example, heavy investments by Americans abroad tend to raise the rate of exchange in New York on London, Paris, and other centers. Gold thus tends to flow out, the rate of discount is raised, credit is checked, and commodity prices tend to fall. But this state of affairs makes the United States a good place in which to buy; and since foreigners are supplied with additional purchasing power, in consequence of the American investments abroad, an export of goods from the

United States will ensue, which, in turn, will tend to reestablish a normal equilibrium of prices, discount rates, and exchange rates. Similarly, if the Agent General boldly transfers the reparations payments (thus transferring purchasing power from Germany to the allied countries) an increased demand for German goods will develop and a surplus of German exports will ensue. Thus it does not seem impossible that a *real* payment of the reparations (that is, in terms of goods) can eventually be effected. There is of course a limit to the amount of surplus exports that a nation can produce; but the Dawes annuities, representing, as they do, only about 5 per cent of the German national income, do not appear to be impossible.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. German labor is likely to feel the weight of the reparation payments for two reasons: first, the taxes imposed are likely to be of such a nature (in part at least) that the cost of living will be raised; second, the fall in the German mark (due to the sale of the marks paid to the Agent General) will tend to stimulate exports, leaving a smaller volume of goods in Germany, thus raising prices without at the same time raising money wages. Discuss the likelihood that the reparation payments will give rise to much industrial unrest.

2. The indemnity imposed upon France, following the Franco-Prussian War, was 5,000,000,000 francs. French investors who held foreign securities were induced to exchange these for French government bonds. These foreign securities were sold in other countries and from the proceeds of this sale a sufficient volume of foreign exchange became available to pay the indemnity.

What bearing does this statement have upon the argument so frequently made that an indemnity can only be paid in exports of merchandise?

3. Up to the end of 1927 about half the German reparation payments have been made in deliveries in kind, particularly to France and Belgium. These deliveries have consisted of machinery, fertilizers, paper, coal, and coke. In what respect, if any, is this form of payment superior to payments made in the form of foreign exchange?

CHAPTER XXXVI

THE LAW OF COMPARATIVE ADVANTAGE

It is the purpose of the present chapter to explain the fundamental principles that determine the currents of international trade. We have postponed this chapter until the principles underlying the distribution of a nation's income into wages, rent, interest, and profits had been dealt with, because this analysis is necessary in order to understand thoroughly the law of comparative advantage and its effect on international trade.

In this chapter we shall consider (1) why nations engage in international trade, (2) the extent to which it is advantageous for a nation to engage in international trade, and (3) the grounds for the frequently expressed fear of foreign competition.

The Basic Reason for International Trade. The basic reason for trade of any sort, whether domestic or international, is the advantages to be derived from industrial and geographical specialization. Nearly everyone in modern industrial societies produces for the market. Such specialization involves, of course, an enormous amount of trade; but this trading is worth while, for the specialization that it makes possible increases our total output enormously. Trade and transportation are in themselves so much economic waste except in so far as they serve to develop division of labor and territorial specialization and thereby increase the total volume of production. Thus it is often said that it is a great waste for farmers to ship their hogs and cattle to Chicago only to have pork and beef shipped back to the rural towns. Or take the example of the Wisconsin cheese factories which sell their output to wholesalers and storage men in Chicago, where the product is cured, graded, and packed. These wholesalers sell to retailers in all parts of the country, including the small Wisconsin towns where the cheese was originally made. This looks like wasted effort, but it is really efficient division of labor. The market is so restricted in these small towns that the work of storage, packing, grading, and wholesaling cannot be efficiently done there either by the retailers as a side issue or by specialized dealers.

Trading not only makes human specialization, or division of labor, possible, but also territorial and geographical specialization. If the population is mobile it will tend to become distributed over the country according to the natural advantages of the sites and resources of the different regions. Each part of the country specializes more or less in a limited number of products. No doubt any one of these regions could, if necessary, produce a much greater variety of products, but it is not economical to do so, since they can be produced to greater advantage elsewhere.

Comparative Cost in Regional Trade. Let us suppose for the sake of simplicity that a nation's consumable goods can be classified under three heads, A, B, and C, and that the country can be divided from the standpoint of the nature of its resources into three regions, Nos. 1, 2, and 3. Let us suppose that any one of the three regions can produce all the three varieties of finished products named, but that region No. 1 has a comparative advantage in the production of A, the next best product being B; that region No. 2 has an advantage in the production of B, the next best product being C; and the region No. 3 can best produce C, the next best product being A. Now any one of three possible situations may exist: (1) we may find that region No. 1 produces A exclusively, that region No. 2 produces B exclusively, and that region No. 3 produces C exclusively; (2) we may find that region No. 1 produces commodities A and B, that region No. 2 produces commodities B and C, and that region No. 3 produces commodities C and A; or (3) we may find each of the three regions producing all the three commodities.

Certain very definite factors, which can best be set forth by the aid of diagrams, will determine which of the three possible contingencies will actually obtain.

Let us suppose that the line *OV* (Fig. 50) represents the marginal productivity[1] of labor[2] in this country. The diminishing productivity of labor in the three regions and for the three different commodities is represented by the curves *PP*. From the diagram it will be seen that OM_1 units of labor are employed in the production of commodity A in region No. 1, that OM_2 units of labor are employed in the production of commodity B in region No. 2, and that OM_3 units of labor are

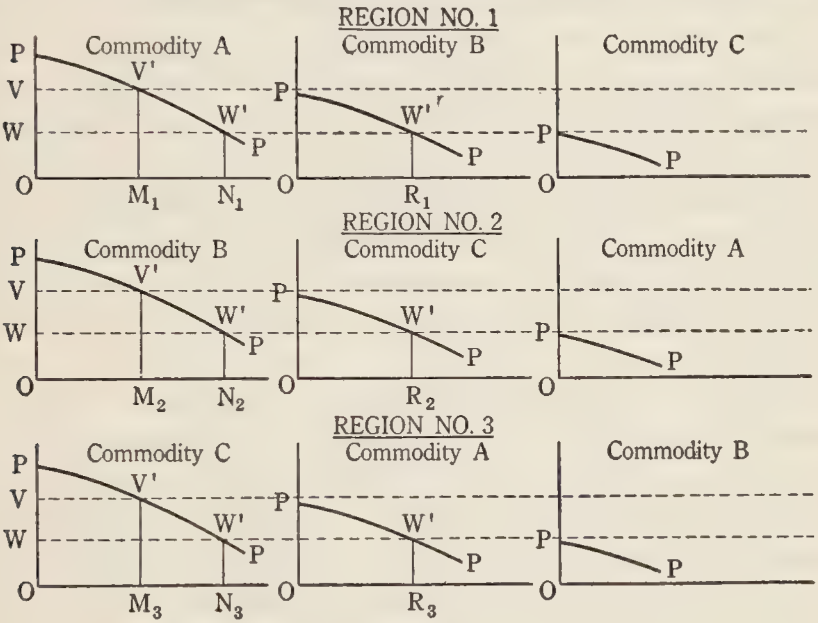

FIG. 50

---

[1] This term is used throughout this chapter to mean *value* productivity.

[2] Attention should be called to the fact that, for the sake of simplicity, we have centered our attention in this chapter exclusively on *labor*. Precisely the same analysis applies, however, to the other factors of production. We have presented the theory of international trade, not in terms of the discredited labor theory of value, but in terms that harmonize with the modern theory of value and the productivity theory of distribution.

employed in the production of commodity C in region No. 3. There is no more labor to be employed. The marginal product of labor is the same (*OV*) in all three regions. This naturally follows if we assume uniformity in skill required and perfect mobility of labor. Now, under the conditions indicated in the diagrams, it is clear that no labor could be employed in the production of commodities B and C in region No. 1. The reason for this is that even under the most favorable combination of factors possible, a unit of labor cannot produce enough of these commodities to yield a return equal to the marginal-value product of labor employed in the production of commodity A. Likewise in region No. 2 the production of commodities A and C, even under the most favorable circumstances, is not sufficient to equal the value product of the marginal laborer employed in the production of commodity B. Similarly, in region No. 3, only commodity C can economically be produced.

But now let us suppose that population increases in all three regions. The margin of production is now pushed lower down in each of the three regions. The marginal productivity of labor is now represented by the line *OW*, and the number of laborers employed in region No. 1 by $ON_1 + OR_1$, the number in region No. 2 by $ON_2 + OR_2$, and the number in region No. 3 by $ON_3 + OR_3$. It will now be clear from an examination of the charts that region No. 1 will produce commodities A and B, that region No. 2 will produce commodities B and C, and that region No. 3 will produce commodities C and A. In region No. 1, for example, where in the former illustration it was uneconomical to produce commodity B, it is now possible to do so because of the fact that the margin of productivity has fallen to *OW*. A unit of OR_1 labor producing commodity B will now yield the same value product as a unit of ON_1 labor engaged in the production of commodity A. Similarly, in region No. 2 the marginal-value product of OR_2 labor producing commodity C is equal to the marginal-value product of ON_2 labor engaged in the production of commodity B.

While under the lower margin of productivity each region would be engaged in the production of two commodities instead of one, it should be noted, however, that each region is still specializing largely in the commodity in which it has the greatest advantage. For example, region No. 1 produces much more of commodity A than of commodity B, for it has a greater advantage in the production of A; and region No. 2 produces a larger quantity of commodity B than region No. 1, for it has an advantage over region No. 1 in the production of this commodity.

From the principles which have so far been developed it should be clear that if the population continued to increase and the marginal productivity of labor were pushed down still farther, the time would come when each region would produce all three commodities, though relatively more of the commodity in which it has an advantage. It would not have been possible, however, for each region to produce all three commodities were it not assumed that each region had a considerable degree of diversification of natural resources. The more limited a region is with respect to the versatility of its resources and its people, the more specialized it will be and the fewer products it will produce. It also follows from the principles illustrated above that the undeveloped region will (if it has trade connections with the rest of the world) be more highly specialized than it will be later in its history when it becomes more densely populated. In a new country people naturally pursue those occupations and industries that give the greatest return. Gradually less advantageous industries are developed.

There are two factors that chiefly determine the extent to which any region will engage in diversified production: (1) the variety of its natural resources and the versatility of its people; (2) the margin of productivity, which depends, so far as the human factors are concerned, (*a*) upon the size of the population relatively to the extent and richness of the natural resources and the abundance and efficiency of capital equipment.

and (*b*) upon the training and technical efficiency of the population itself including both manual and mental workers.

Comparative Cost in International Trade. We have purposely assumed that all three regions were a part of the same country and that the population was perfectly mobile. As a result the marginal product of labor is identical in all three regions. We have done this to simplify the problem and develop certain fundamental principles which apply particularly to domestic trade and commerce but which also serve as foundation principles for the study of the law of comparative advantage in international trade. From the standpoint of the present discussion the essential difference between international trade and domestic trade is this: in international trade there are
1 national barriers to mobility of population — some rooted deep in the traditions of race, language, and country, and others resting on legislation, such as laws restricting the free flow of immigration. Moreover, the personal efficiency of
2 labor in one nation may be much higher than in another nation. As a result of the immobility of population and the differences of personal efficiency the marginal productivity of labor varies very greatly from one country to another.[1] Consequently many countries now produce commodities which they could not economically produce were the marginal productivity of labor uniform throughout the modern industrial world.

Let us take, by way of illustration, two countries, K and L. These two countries, let us assume, are both similarly endowed by nature with a great variety of natural resources and with equally efficient populations. But country K has a high margin of productivity and country L a low margin. The necessary principles can be illustrated more clearly by reference to Fig. 51.

In these diagrams it is assumed that both countries have exactly the same natural advantage in the production of each of the four commodities. The same set of productivity curves will therefore serve for both countries. But the population of L is very much greater than the population of K, and as a result the natural resources of L are utilized much more intensively than in K. Consequently the marginal productivity of labor is much lower in country L than in K. Thus in the diagram, *OL* represents the marginal productivity of labor in country L, and *OK* the marginal productivity of labor in country K. As a result, there will be a very great difference in the relative distribution of the gainfully employed persons engaged in the production of the four commodities in these countries in spite of the fact that the natural advantages of both countries are identical. Thus *OM* persons will be employed in the production of commodity A in country K, whereas *ON* persons will be so engaged in country L; *OR* persons will be producing commodity B in country K, whereas *OS* persons will be so engaged in country L; *OT* persons will produce commodity C in country K, whereas in country L the number so occupied will be *OU*; in country K none will be engaged in the production of commodity D, whereas *OZ* persons will be so engaged in country L. In spite of the fact that the natural advantages of both countries are identical, it is not economically possible in country K to employ labor in the production of commodity D, because the marginal productivity of labor in the case of commodities A, B, and C is too high to

---

[1] Attention should, however, be called to the fact that there is actually no such perfect mobility of labor in any country as we have assumed in this argument. In the United States, for example, there are very wide differences in wages in different sections of the country.

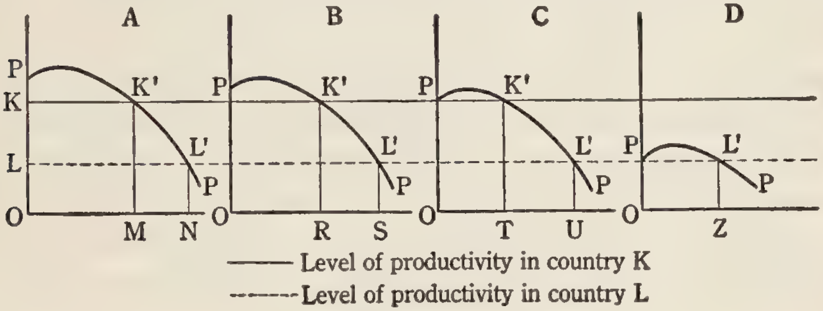
A
B
C
D
P
K
L
O
K'
L'
P
M
N
R
S
T
U
Z
—— Level of productivity in country K
------ Level of productivity in country L

FIG. 51

permit such employment; in other words, the alternative opportunities for the employment of labor in the production of A, B, and C are so good that no one can afford to employ labor in the production of commodity D. On the other hand, in country L the alternative opportunities in the employment of labor in the production of commodities A, B, and C are so poor or, in other words, the marginal productivity of labor is so low that it is economically feasible to employ labor in the production of commodity D. Business men would say that it is impossible for country K to compete with country L in the production of commodity D, because of the fact that wages are higher in K than in L. This statement is perfectly true; but what is often overlooked is the fact that wages in country K are high merely because the alternative opportunities are so good that employers can pay these high wages in the production of commodities A, B, and C and still make a competitive profit.

A study of the foregoing diagram should make it clear that a considerable amount of trade would be likely to take place between these two countries in spite of the fact that both were presumed to have identical natural advantages. Certainly country K will import her entire supply of commodity D. Very likely country K would import a part of her supply of commodity C, since the production of this commodity is very limited in country K and very extensive in country L. There probably would be little if any trade in commodity B, since the relative production of this commodity in the two countries corresponds roughly to their respective populations. Probably country K will export considerable quantities of commodity A to country L, since the production of this commodity in K is relatively large for her population.

N.B.

From this illustration an important principle of international trade may be derived. The extent of the trade and commerce between two nations depends in part upon differences in the marginal productivity of the two countries. Since wages depend upon the marginal productivity of labor, this is merely another way of saying that the extent of trade between two countries depends in part upon the relative wage levels of the two countries. The wage level in a country may be so high that certain industries cannot be developed economically. In such a country only the industries with superior natural advantages will flourish. In these industries the productivity is so great that in spite of high wages the *labor cost* is no higher than in low-wage countries. In low-wage countries, however, even the industries with low natural advantages can be profitably carried on, for in spite of the low productivity labor costs are not excessive, because of the low wage level.

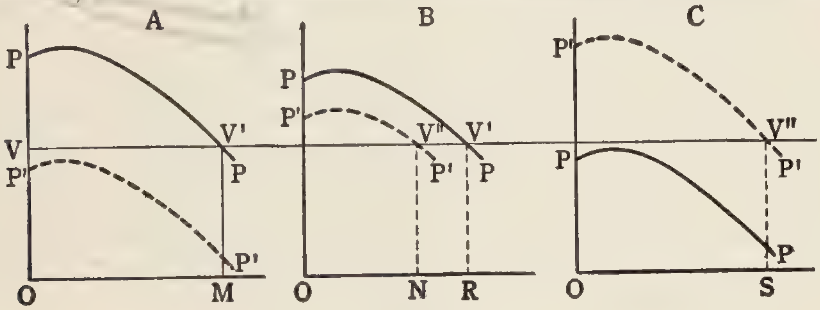

FIG. 52

The second important reason for international trade is to be found in the advantages of one country over another in the production of certain commodities. To illustrate this principle clearly we shall assume that the wage levels of two countries are identical, but that in the production of certain commodities advantages are enjoyed by each country over the other. These advantages may be due to differences in natural resources, nearness to markets, special training and skill of the population, or the efficiency of capital equipment. The principles involved are illustrated in Fig. 52.

The diminishing-productivity curves for country X are represented by the lines *PP*, and for country Y by the lines *P'P'*. The marginal productivity of labor in both countries is represented by the line *OV*. The number of laborers employed on the several commodities in country X is represented by *OM* and *OR*, and the number of laborers employed in country Y is represented by *ON* and *OS*. It is clear from the diagrams that country X has a great advantage over country Y in the production of commodity A, that its advantage is much less in the case of commodity B, and that in the case of commodity C, country Y has a very great advantage over country X. All of X's gainfully employed population is engaged in the production of A and B, and none of its labor is employed in the production of C, since it can be more advantageously employed elsewhere. On the other hand, none of the gainfully employed population of country Y is engaged in producing A, for its labor can be more advantageously employed in producing B and C. Both countries produce B, but in each country the number of laborers so employed is less than the number engaged in the production of the commodity in which that country has the greatest advantage. It should also be noted that while both countries produce commodity B, country X has a larger number of laborers employed in the production of this commodity than country Y, because of the fact that the former country has an advantage.

In this illustration it is clear that X will export commodity A to Y and that Y will export commodity C to X. Both countries produce B, but it is probable that Y does not produce enough to satisfy its market and therefore imports a portion of its supply from X.

These diagrams illustrate the principle that even though the productivity and wage levels are identical in two countries, trade will normally take place between the two countries because of the advantage that one country has over another in the production of certain commodities.

Both the illustrations which we have used are highly artificial. Nowhere do we find two countries with exactly the same productivity and wage levels, and nowhere do we find two countries which have exactly the same natural advantages. Countries, in fact, differ greatly, not only in the level of productivity (upon which the wage level depends) but also in the natural advantages or disadvantages in the production of various commodities. Let us take Great Britain and the United States by way of illustration. The United States has a great natural advantage in the production of wheat and coal and, to a less extent, in the production of cloth. Let us assume that Fig. 53 illustrates the natural superiority of the United States in the production of these commodities.

The diminishing-productivity curves for the United States are represented by *PP*, and for Great Britain by *P′P′*. The numbers gainfully employed in the production of the various commodities in the United States are represented by *ON*, *OS*, and *OT*, and for Great Britain by *OM*, *OR*, *OU*, and *OZ*. The

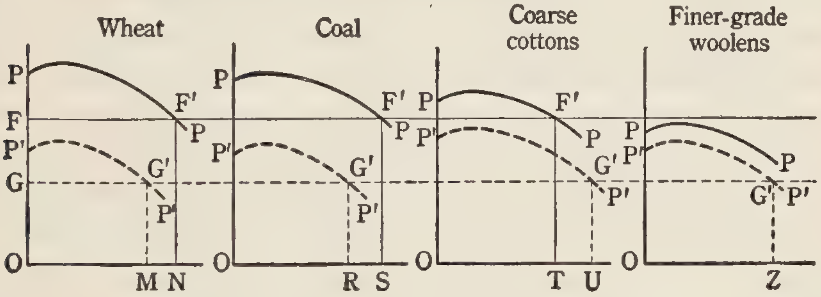

Fig. 53

marginal productivity (and hence the wage level) for the United States is represented by *OF*, and for Great Britain by *OG*. From the productivity curves it is clear that the United States has a natural advantage over Great Britain in the production of every one of these commodities; but the advantage is greatest in the production of wheat, second in the production of coal, third in the production of the coarser cottons, and least in the case of the finer-grade woolens.

Since the United States has a natural superiority in the production of all commodities, it may at first thought appear that it could not profitably import anything from Great Britain. It is, of course, easy to see that if Great Britain has a natural superiority to the United States in the production of certain goods it would be the normal thing for us to import those goods while we in turn would export to Great

Britain the goods in which we have an advantage. But it is not so easy to see that even though we have a natural superiority over Great Britain in every line of production it will still be advantageous for us to import many things from that country. The United States cannot afford to produce everything in which it has an advantage over Great Britain.

A study of the curves will show that Great Britain will produce every one of the commodities named in spite of the fact that she is at a natural disadvantage in every one. On the other hand, the United States will produce only the three commodities in which it has the greatest advantage in spite of the fact that it has a natural advantage in the production of all four commodities. It cannot afford to produce the finer grade of woolens at all, because the marginal productivity of labor in the production of wheat, coal, and the coarser cottons is so high that labor and capital are attracted to these fields. Labor cannot produce a value product of fine-grade woolens sufficient to equal the marginal product in the other fields. In other words, the wages that can be earned in the production of wheat, coal, and cottons are so high that no employer can afford to employ labor in the production of the finer grade of woolens.

Let us analyze this problem a little more closely. Assume that the daily wage in the United States is $4 and that it is $2 in Great Britain. Suppose that an American workman can, under the most favorable combination of factors, produce $3.75 worth of woolen cloth, and that a British workman can, under the most favorable combination of factors, produce $3.50 worth. Since American workmen can earn $4 in other lines of production, it follows that they cannot profitably be employed in the production of woolens. On the other hand, since British workmen can earn only $2 in other lines, it follows that they can profitably be employed in the production of woolens, and labor and capital will flow into this industry until the marginal product of labor is reduced to the general level of $2.

It should now be clear that the United States will import its total supply of fine-grade woolens in spite of the fact that an American workman can produce more yards of woolen cloth a day than a British workman can. On the other hand, although Great Britain will produce a part of her wheat supply, she will import most of it, since, except for the land most favorable to wheat production, her labor can be employed more profitably elsewhere. In the case of the coarser cottons the United States imports a considerable amount in spite of the fact that it has a natural advantage over Great Britain. A part of its domestic market will be supplied by its own labor — that part which can be produced under sufficiently favorable conditions to pay the wages which can be earned elsewhere.

We may now summarize the more important conclusions as follows:

1. A nation's labor and capital will naturally flow into those occupations which will yield the *highest* returns.

2. From the principle given above it follows as a corollary that a country may import its total supply of certain commodities even though it has a natural superiority over the competing country. The reason is that its population is fully employed in the production of goods in which it has a still greater advantage.

3. Even though a country is inferior in the production of many or even all commodities, it may still be able not only to supply its domestic market, but even to export certain commodities to the superior country. The natural inferiority in production will, indeed, mean a low wage level, but it will not prevent its capital and labor from being employed. An inferior nation need therefore have no fear that other nations with superior natural advantages will be able to undersell it in all lines and thus drive it out of business. The superiority of the foreign nations will force their employers to pay high wages, whereas the inferiority of the country in question will compel its employers to pay low wages. Thus the wage level of the superior country will be so high that the inferior country with its low wage level can successfully compete in many lines.

4. On the other hand, a country with a high wage level need have no fear that low-wage countries will be able to undersell it in all lines of business. It must not be forgotten that such a country enjoys a

high wage level *because* of its superiority in production. Therefore, even with the high wages its *labor costs* are no higher than in low-wage countries, and in many lines probably not as high. In this connection it is *labor costs*, not *wages*, that count. What is often forgotten is the fact that wages are high simply because the country does enjoy a superiority in production. If it were really true that low-wage countries could undersell the high-wage countries in all lines, then, of course, it would be impossible for the high wages in the superior country to continue at their former high level, and the inevitable result would be a readjustment. But the fact that the high wage level continues is itself proof conclusive that the industries of the country can pay it.

The Menace of Foreign Competition. Why, then, do we hear so much in the business world about the fear of foreign competition if in fact these things automatically adjust themselves? The answer is that it is through the competition of entrepreneurs that the adjustments take place. It is this competition that forces wages up or down, as the case may be, and into line with the marginal productivity of labor. The readjustments are constantly forcing certain entrepreneurs to the wall — entrepreneurs who find themselves unable to pay the going rate of wages and to sell their products at the market prices. The pressure of competition is a real one; and it is no wonder that employers are much concerned over the wages paid by their competitors, both foreign and domestic, as well as the prices at which they sell their products.

Moreover, situations sometimes arise under which foreign competition becomes exceptionally serious. If the wage level has by extraordinary circumstances been raised in one country or lowered in another above or below the normal level established by the marginal product of labor, it is clear that the low-wage country will prove a dangerous competitor to the high-wage country. Under these circumstances the country with a wage level lower than the marginal productivity of labor would warrant has low labor costs, whereas the country with an abnormally high wage level has exceptionally high labor costs. These conditions would permit the country with low labor costs to undersell the country with high labor costs. Yet this situation would set forces in motion that would tend to reëstablish in each country a wage level that would correspond to the marginal product, for the country with low labor costs would capture the market and expand production, whereas the country with high labor costs, unable to get the orders, would be forced to close its shops and turn its wage-earners out on the streets. Expanding production in the former country would create a scarcity of labor, and the resulting competition for workers would gradually raise the wage level to the marginal-productivity level. On the other hand, hard times and unemployment in the high-wage country would exert a powerful pressure in the direction of a reduction of wages.

Many illustrations of such maladjustments as we have just spoken of could be cited. Often they constitute problems of major importance in modern life. Take the case of England and Germany in the post-war period. With the drop in prices in 1920 British wages failed to decline proportionately, partly because of the inertia of wages as compared to wholesale prices, partly because of social legislation, and partly because of the organized effort of the trade unions. Consequently wages pressed severely on profits, and large numbers of workmen were unemployed. On the other hand, inflation continued in Germany until near the close of the year 1923. During this period of terrific inflation German wages lagged far behind prices. As a result of these opposite price movements and the wage lag, British wages were abnormally high and German wages were abnormally low — abnormal, that is to say, in view of the marginal productivity of labor in each country. Labor costs were high in England and low in Germany; therefore England found it difficult to compete with Germany in the foreign markets.

Similarly, if minimum-wage laws, social insurance, and collective bargains with trade unions have the effect of raising labor costs, it is clear that the entrepreneurs in countries with lower labor standards have an advantage.[1] A larger slice of the market will consequently go to the country with lower labor costs, and the country with superior labor standards can retain them only at the expense of a certain amount of unemployment. In other words, the high labor cost will tend to force a certain number of marginal firms out of business, leaving the field to the more ably managed firms or to those having superior sites and resources. Hence countries with advanced labor standards, such as England, are necessarily interested in raising labor standards in less advanced countries in order to protect their own standards. If the standards are raised in all competing countries, a steady pressure is applied on the entrepreneurs of all countries to raise the level of productivity; but if one country alone is seeking to maintain high standards, it is likely to be worsted in the open market by the unscrupulous low-standard countries. Not infrequently a situation of this sort is found between different regions within the same country. Thus, in the United States the nonunion coal mines in West Virginia are able to reduce wages in periods of depression, whereas the union coal fields, working under two-year or three-year contracts, are unable to do so. Thus in the period of depression when the market is limited the nonunion mines are able to undersell all but the better union mines. This forces the union miners to pay for their superior wages and working conditions by bearing the brunt of unemployment when the trade is depressed. This indicates the danger in raising standards higher than is justified by the industrial situation. On the other hand, if the pressure of high labor standards is removed, entrepreneurs are likely to seek to reduce costs by lowering wages rather than by increasing the productivity of the industry through the elimination of waste and improved management or technical processes. The competition is clearly unfair when high standards are exacted in certain regions and not in others.

---

[1] It should not be forgotten that these measures do not of necessity raise labor costs; in fact they have often lowered the labor cost through the increased productivity stimulated by establishing better conditions for the wage-earners. The increase in productivity may be due either to increased efficiency of the workmen themselves or to improved processes and methods introduced by the firm.

Another matter that deserves brief consideration in this connection is the effect of the unequal development of economic efficiency in different countries. Before the war Germany was making exceptionally rapid strides in industrial progress. Was her growing prosperity an economic advantage or disadvantage to other countries — to England, for example? That depends. In so far as Germany's increased productivity occurred in industries in which British labor and capital were chiefly engaged, England would suffer. The enlarged output from these industries would lower the exchange value of their products, thus lowering the value product of British industries, even though the physical product was as high as before. With lower marginal productivity real wages would tend to fall until labor costs for British entrepreneurs reached the level of German labor costs. This would be the price that England would have to pay to hold her place in competition with Germany. In so far as Germany's progress developed in lines in which England was predominantly a consumer rather than producer, England would gain, since the increased output of these goods would raise the exchange value of British industry. It is often true that a country gains from the increased productivity of her neighbors, but not always. If she is a consumer of the goods in question she gains; if she is a producer of similar goods, she may be a loser unless she can readjust her production successfully on the new basis. The doctrine of harmony of interests between nations has its limits.

It is sometimes said that a country like England will suffer in so far as Australia, India, China, Russia, and other countries develop industrially and become more self-sufficient. This is not necessarily true. It is not unlikely that as these countries expand their production their increased purchasing power will improve England's foreign market. Even though these countries do furnish in part the same types of goods and services which England offers, nevertheless they may produce so much more of other sorts of goods that, relatively, England's products will be scarcer (and therefore more valuable) than now. On the other hand, it is possible that these undeveloped countries may eventually produce so much of the goods and services which England (both because of her resources and her people) is best adapted to furnish that her position will be weakened. The value product of a nation (and therefore a nation's income) depends in the modern order not only on its productive efficiency (and this in turn rests on (1) its natural resources, (2) its capital equipment, and (3) the efficiency and genius of its people) but also in large part on what is produced elsewhere. Any of these factors will affect England's future position in world trade.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. "The only reason why we gain from international trade is that the national resources and abilities of the people are different in the different countries." Do you agree? Why, or why not.

2. "The fact that the wage level is not the same in different countries is one of the leading causes of international trade." Discuss.

3. "Under free trade, goods will tend to be shipped from a low-wage country to a high-wage country until wages are equalized." Show that this is not true, and point out the reasons why.

4. "If a country has a natural or technical superiority over all other countries, it cannot afford to import that commodity." Do you agree? Why, or why not.

5. "The high wage level in Australia prevents her from developing many industries that might flourish there." Show that this is true, and point out why the wages are high. Does it follow that it is desirable for a country to have a low wage level?

6. "High wages make labor cost high in some industries but not in others." Why? Under the circumstances what will become of the high-labor-cost industries?

7. "It is an old fallacy that a nation suffers from the prosperity of its neighbors." Does a nation always profit by the prosperity of its neighbors? Does it sometimes profit? Explain.

8. "It is important that wages and working conditions be made uniform throughout the world; otherwise labor costs will be high in some countries and low in others, and this results in unequal competition." Do you agree? What fundamental point is overlooked? If wages are equal the world over, are labor costs likely to be equal? Explain.

9. "It is difficult for a nation to improve its labor standards by legislation so long as other countries do not do so." Show that this is true, and explain the reason why.

10. "During the war many of the industrially backward countries, such as South Africa, Australia, India, and China, were deprived of part of their customary supply of manufactured goods from Europe. Thus the war stimulated the development of manufacturing in these countries. After the war the doors were closed to the European products, and this accounts, in large part, for the post-war European depression." Discuss.

CHAPTER XXXVII

TARIFF AND FREE-TRADE PRINCIPLES

Until recently duties on imported articles constituted an important source of the revenue of our Federal government. However, since the import duties yield at the present time only about 15 per cent of the Federal revenue, it will be seen that the tariff as a producer of revenue is no longer of outstanding significance. The other function of the tariff — to protect domestic producers against foreign competition — continues to occupy a place of great importance in modern governmental policies.

In the early history of the American union it was certainly true that the tariff was of greater significance as a producer of revenue than as a device to protect American manufacturers against foreign competition. Not until the close of the Napoleonic wars and the War of 1812 was the tariff seriously made use of as a protective instrument. During the course of these wars commerce between the neutral states and England had been largely interrupted as a result of the restrictions imposed on international trade by the Berlin and Milan decrees issued by Napoleon, the retaliatory orders in council issued by the British government, and finally by the embargo acts of the United States government. This obstruction to commerce forced the Americans to rely on themselves for the supply of manufactured products to a far greater extent than before. When the war was over and the English manufacturers began to dump their surplus products on the American markets, it was inevitable that the owners of the newly formed and weak American industries should demand protection against the older and more firmly established English firms. In response to this demand the tariff act of 1816 was passed. In particular this act brought a considerable measure of protection to the rising textile and iron industries — the two industries that first developed the factory system in the United States. Still higher protection was granted in the acts of 1824 and 1828. In the act of 1828 an appeal was made to the farmers of the new West to lend their support to the protective system by including among the protected articles wool, hemp, and flax.

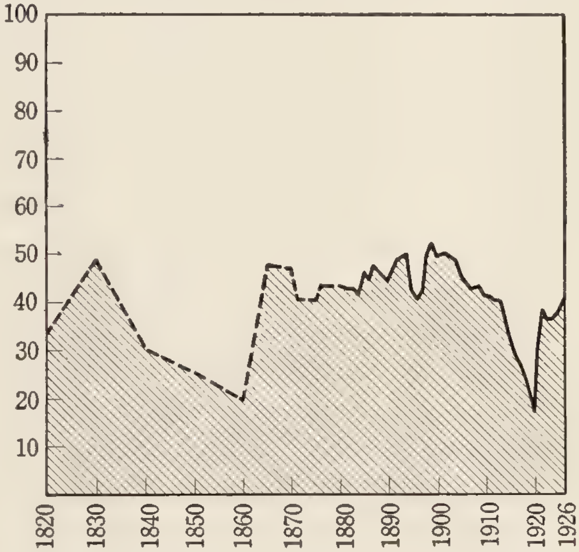

FIG. 54. Ratio of duties to total dutiable imports entered for consumption. From the Statistical Abstract, 1926

In the act of 1833 the country began, under the leadership of the agricultural South, which exported its surplus product of cotton and tobacco and therefore needed no protection, to move in the direction of freer trade. This compromise act provided that the duties should be progressively reduced until 1842. But in 1842, on the return to power of the Whigs, higher rates were again imposed. The trend toward tariff reduction returned, however, with the act of 1846, and this was followed in turn by a still lower tariff in the act of 1857, when the rates of import duties to the total value of dutiable imports fell to the level of approximately 20 per cent.

In 1861 the Republican party, which favored the policy of protection, assumed control of the government. Three tariff acts were passed during the Civil War period, each higher than the preceding one, until in the act of 1864 the level of duties was raised to nearly 50 per cent. Then began a long period of high protection during which the ratio of duties to taxable imports remained at the high level of from 40 to 50 per cent. Not until the Underwood tariff of the Wilson administration in 1913 was any material reduction achieved. Under this act the ratio of duties to taxable imports fell to about 20 per cent in 1918–1920. With the return of the Republicans to power in 1921 a temporary tariff measure of a highly protectionist nature was passed, and this was followed by the act of 1922, which restored the tariff wall to approximately the 40 per cent level.

The average height of our tariff wall from 1820 to 1926 can be seen at a glance from Fig. 54. Except for the two decades before the Civil War and the period of the Wilson administration the tariff wall (the ratio of duties to taxable imports) has persistently stood at a level of from 40 to 50 per cent.

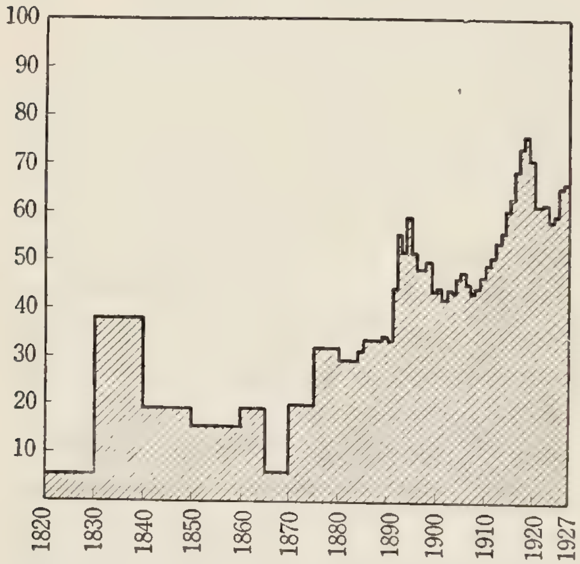

FIG. 55. Per cent of value of imports admitted free (1821–1870, per cent of total imports; 1871–1926, per cent of total imports less reëxports). From the Statistical Abstract, 1926

On the other hand, if we consider (see Fig. 55) the percentage of the total value of our imports which is free of duty, the tendency has been almost steadily in the direction of freer trade. In the decade of the twenties (nineteenth century) only 6 per cent of our imports (measured in terms of value) were duty-free. In the thirties the percentage rose to nearly 40 per cent, but from 1840 to 1865 it fell to less than 20 per cent. In the five-year period following the Civil War only 7 per cent of our imports, by value, were on the free list. From this time on the proportion rose to nearly 60 per cent in 1874, then declined to 42 per cent in 1901, after which it rose to 75 per cent in 1918, the highest point ever reached. Under the Republican tariff acts of 1921 and 1922 the percentage has ranged from 60 to 65 per cent. Thus even under the present protective system nearly two thirds of our imports are duty-free.

A much larger proportion of our finished manufactured goods are in the free list now than formerly. In 1926 one third, by value, of the imports of this class of goods were free of duty, but it is especially crude and semimanufactured materials that are admitted without tariff restrictions. Among the important commodities on the free list in the tariff act of 1922 are crude rubber, raw silk, coffee, hides and skins, petroleum, tin, and copper. These seven commodities alone constitute over one third of our total imports.[1] Table XXXVII, on page 704, gives the trend toward free-trade imports by broad classes of commodities.

Arguments in Support of the Protective Tariff. *The Infant-Industry Argument.* The arguments that have been advanced in support of the tariff have shifted from time to time in our history. Changing economic conditions have altered the arguments brought forward, and each argument advanced has been suited to the peculiar conditions of the period in question. The earliest theory advanced in support of the protective

[1] LEADING IMPORTS INTO THE UNITED STATES IN 1926

	VALUE	PERCENTAGE OF TOTAL IMPORTS
Five leading imports		
Rubber (crude)	$506,000,000	11.4
Silk (raw)	393,000,000	8.9
Coffee	323,000,000	7.3
Paper and wood pulp	261,000,000	5.8
Sugar (cane)	233,000,000	5.2
Animal products		
Furs and fur manufactures	117,000,000	2.6
Wool and mohair	107,000,000	2.4
Hides and skins	97,000,000	2.2
Mineral products		
Petroleum and petroleum products	125,000,000	2.8
Tin and tin products	105,000,000	2.4
Copper (ore and manufactures)	100,000,000	2.3

TABLE XXXVII. PERCENTAGE OF IMPORTS (MEASURED BY VALUE) FREE OF DUTY

	1821	1860	1900	1926
Crude materials	58.1	42.7	75.6	82.8
Crude foodstuffs and food animals		92.1	70.2	81.2
Manufactured foodstuffs		7.4	17.4	13.4
Semimanufactured goods . . .	12.1	18.4	34.0	70.7
Finished manufactured goods . .	0.2	1.9	8.4	34.8

tariff was the infant-industry argument. According to this theory a protective tariff policy is especially valuable in a new country where the resources and people are well adapted for the industries which it is proposed to protect, but these industries find it difficult at first to compete with the industries which are already firmly established in the older countries. The protective tariff helps to get these industries on their feet; and it is believed that once they have become going concerns they will be able to face world competition without any sheltering governmental protection. The argument was effectively made use of after the Napoleonic wars, when our infant manufacturing industries found it difficult to face the competition of the better-established English manufacturers.

If a protective tariff is to be successfully maintained politically, the system must be made to appeal to the general public and not merely to a specially favored class of producers; or at least it must be shown that the benefits accruing to a special class will eventually accrue directly and indirectly to other classes as well. The infant-industry argument sought to show that the building up of these industries would be beneficial to the entire nation, since it would eventually enable the domestic consumers to secure the products of these industries on more favorable terms than from foreigners. Once the industries in question were firmly established, the whole nation, it was argued, would benefit.

The Home-Market Argument. Historically the home-market argument came next. The infant-industry argument was not sufficient to satisfy the demands of the protectionists, since it could be made to apply only to those industries which the resources of the country and the abilities of its population specially favored. It was next the desire of the protectionists to show that it was economical to develop even those industries which could never hope to face foreign competition on a free-trade basis. The home-market argument was admirably suited to this purpose.

This argument was directed in particular to the farmers, who, after the Napoleonic wars, were suffering from inadequate foreign markets for their products. Farmers have always been hard hit by post-war deflation of prices, and the post-Napoleonic agricultural depression had many points of similarity to the recent agricultural depression which followed the World War. Under the leadership of Henry Clay the protectionists now came forward with what appeared (superficially at least) to be a solution of the problem. And this was the proposal: American manufactures were to be built up by means of a protective tariff. The development of these industries would create a large urban population which would furnish a good domestic market for farm products. With a larger proportion of the population engaged in manufacturing, a more favorable balance would exist between agriculture and industry. Thus it was argued that a tariff which directly protected manufacturing was indirectly beneficial to agriculture as well.

The High-Wages Argument. Until the twenties (nineteenth century), during which decade the wage-earners in many states gained the suffrage, it had been urged that a protective tariff was necessary because of the high level of wages in America compared with the wages paid in Europe. Manufacturing could not hope to develop, it was argued, in a country that was faced in its foreign competition with the handicap of extraordinarily high wages. But now that the wage-earners had become voters it became necessary to gain their support for the cause of protection. So the argument was shifted, and it was now made to appear that the protective tariff was the *cause* of high wages, and that a protective system was necessary to *maintain* high wages. Without this protection foreigners would be able to undersell American producers, and this in turn would either drive the latter out of business or compel them to lower wages in order to compete with the foreign-made goods. Protection would bring increased employment at high wages. Such was the argument advanced by the politicians in the campaigns of 1828, 1832, 1836, and 1840. The tariff was necessary to protect American labor against the pauper labor of Europe. And this argument has been repeated down to the present day.

The National-Defense Argument. Finally the protective tariff was urged as a means of safeguarding the interests of a nation in time of war. Wars disrupt commerce between nations and force each country to rely more or less exclusively upon its own production. It is argued that in a world which is still subject to important international conflicts every generation or so, it is not wise to depend too much upon foreigners for essential products. A certain measure of national self-sufficiency is desirable and even necessary, it is said, in a world which has not yet been able to achieve lasting international peace. This argument does not rest upon economic grounds at all, but purely upon patriotic and nationalistic grounds.

The Validity of the Arguments for Protection Examined. Let us now consider briefly the validity of these arguments. The infant-industry argument is theoretically entirely sound, but practically there is great danger in it. Industries of this character are never found willing to admit that they have arrived at manhood; they are never willing to test their strength by facing foreign competition in a free market. This has been our experience with the cotton, woolen, silk, and iron and steel industries. Having grown up under the sheltering support of the tariff they persist in clamoring for a continuance of its protection. There are, indeed, sound theoretical arguments for protection, as the British economist Sidgwick once said, but in practice free trade is safer.

There is also much sound sense in the national-defense argument, but, as we have already indicated, the argument does not rest on economic grounds. Even though it were detrimental to a nation economically to build up a protective system, it might still be worth while, as Adam Smith pointed out a hundred and fifty years ago, to maintain such a policy from the standpoint of national safety. It is not economical to maintain a standing army, but it may still be necessary to do so, and the same argument may apply to the tariff.

But even the nationalistic argument is not wholly faultless. Tariffs and wars are interrelated in a vicious circle. Tariffs may be necessary as a protective device in case of wars, but the spirit of hostility toward foreign countries is fostered and inflamed by the erection of tariff barriers. A free-trade policy the world over would go far toward establishing international goodwill.

The home-market argument is essentially unsound except in so far as it is linked with the nationalistic argument. The home-market argument overlooks several essential facts. If indeed there is a maladjustment, a lack of balance between agriculture and manufacturing to the disadvantage of agriculture, it follows that labor and capital will gradually be diverted from agriculture into manufacturing, even though no protection were given to manufacturing. Under a protective system this movement would doubtless be more rapid; but it might easily go too far, diverting labor and capital away from farms which from the standpoint of the best economical utilization of the nation's resources ought not to be abandoned.

If, on the other hand, the protection to manufacturing results in a diversion of labor or capital or both from foreign countries to the protected country,[1] the result varies with the particular circumstances. If the immigration of foreign labor is stimulated by the tariff policy (which is doubtful), the foreign market for farm products is likely to be reduced at the same time that the home market is increased, though not necessarily to the same extent. It is not impossible, however, that the foreign population may remain substantially as large as would have been true had there been no emigration. Under these circumstances the population of the world would be greater to the extent that emigration was encouraged by the protective tariff, and world-wide increase in population is, of course, favorable to the farming class or, at any rate, to the farm-owning class. Thus a protective policy might actually benefit farmers, if indeed such a policy really had the effect of stimulating the growth of population in the undeveloped country. And probably a rapid expansion of population in a new country with undeveloped natural resources is, all things considered, beneficial to the nation as a whole. It is true that the wage-earning class is not benefited by such a policy; but since in an undeveloped country the independent proprietors, whether farmers, merchants, or manufacturers, are most numerous the general interest is in favor of an increase in population.

---

[1] It is well known that foreign manufacturers frequently "jump" the tariff wall by building factories within the borders of the protected country. Thus Henry Ford has built several plants in Australia and in various European countries. The International Harvester Company and many others have pursued a similar policy.

It is to be seriously doubted, however, that a tariff policy will stimulate the immigration of foreign labor except in the case of skilled labor, which, without the tariff policy, might not find suitable employment. If the population is not materially increased as a result of the protective tariff, the farmers will lose and not gain from this policy. Their home market will then in point of fact not be increased. Instead labor will be drawn from agriculture into industry, and this will raise the money wages paid for farm labor. Moreover, farmers will be compelled to pay higher prices for the manufactured products, since the domestic manufacturers cannot produce as cheaply as the foreign manufacturers. If the domestic manufacturing costs were not in fact higher, no protective tariff would be needed. Thus the farmers' costs are raised all round as a result of the policy of protection. Moreover, if the tariff policy resulted in no material increase in immigration and population, the farmers would not only lose because of higher costs but also because of a restricted market. If we make our own nation more self-sufficient we thereby force foreign nations to become more self-sufficient also. If they cannot sell their products abroad they are compelled to turn their attention to supplying the home market. As a result agricultural production is stimulated in foreign countries, and so the foreign markets for our farmers are injured. Exporting industries regularly find that their foreign markets are curtailed as a result of tariff barriers and the consequent tendency toward national self-sufficiency. The markets for export industries are best under conditions of free trade. Hence the home-market policy fails to benefit the farmer, first because it is not in point of fact likely to develop the home population (and therefore the home market), second, because it reduces the foreign market, and third, because it raises the money costs of farm labor and manufactured goods purchased by the farmer. Thus the industrial tariff policy tends to lower the prices of farm products and on the other hand to raise the farmer's cost of living and expenses of production. And until the United States becomes a net importer of farm products a tariff on *agricultural* products is of no material aid to the farming population, except in the case of relatively unimportant commodities.

Consider now the higher-wages argument. If the industrial tariff does not aid the farmer in a country like the United States, does it not at any rate benefit the wage-earners? In answering this question there are two possible assumptions. First let us assume that the country is largely undeveloped and that the extractive industries — agriculture, mining, and lumbering — are sufficiently superior to manufacturing (considering the resources, population, and development of the country), so that without the artificial support of the tariff the entire labor force would be employed in these three industries. It is assumed, then, that agriculture, mining, and lumbering, being highly productive, are capable of paying such high wages that manufacturing cannot flourish, for it cannot pay

High [illegible] caused by [illegible]

these wages and compete in the markets of the world. Under these circumstances manufacturing cannot develop at all without a protective tariff. Assume, now, that a protective system is introduced. The prices of imported manufactured goods are now increased as a result of the tariff. This is obviously an injury to the producers in agriculture, mining, and lumbering who buy these commodities. But under the stimulus of these high prices it is now possible for manufacturing to pay as high money wages as can be paid in agriculture, mining, and lumbering, and even to pay wages enough higher to draw labor away from these industries. Thus manufacturing develops at the expense of agriculture, lumbering, and mining. Money wages are raised, but what about real wages? The tariff policy raises the cost of manufactured goods,[1] and in so far as these come within the scope of the wage-earner's budget his cost of living is increased. Will the cost of living rise more than money wages? The answer is unequivocally Yes, because at bottom real wages depend upon the value productivity of the marginal laborer, and this is lowered by the tariff policy. The proof of this statement is not difficult. The real wages of labor in any nation depend finally upon the *quantity* of goods and services which can be purchased with the marginal product of labor in any given industry. The *quantity* of goods and services depends upon the physical productivity in all the various industries taken as a whole. Now will the physical productivity in industry as a whole be greater under conditions of free trade or under conditions of a protective tariff? Surely, under free trade. The physical quantity of goods and services which can be bought with the physical quantity produced by the marginal laborer in manufacturing is not equal, under free-trade conditions, to the physical quantities of these goods and services which may be purchased with the physical product of the marginal laborer in agriculture, mining, or lumbering. When labor engaged in these industries is artificially diverted to manufacturing, a loss in the purchasing power of a day's labor is sustained. By producing exclusively mineral, lumber, and farm products and buying manufactured products abroad the total available quantity of all sorts of goods will be greater than is the case when the imports of manufactured goods are prohibited and labor which is effectively employed in mining, lumbering, and agriculture has to be transferred to the less effective industry; namely, manufacturing. The physical productivity of labor would be lower under a protective-tariff policy, and so the physical quantity of goods of all sorts that can be purchased with the products of the marginal labor in any industry will be reduced. Thus the real wages of labor are lowered as a result of the tariff policy.

---

[1] The prices of the protected goods, of course, tend to rise in relation to other commodities. Moreover, the general price level tends to be high in a high-tariff country. The reason is as follows: when the tariff is raised, imports are curtailed, but exports continue at the old level, for a time. The excess of exports lowers the rate of exchange in the high-tariff country and causes an importation of gold. This additional supply of gold tends to increase the volume of credit in the protected country and therefore to raise prices. Higher prices check a further excess of exports and so operate to check any further inflow of gold. The protected country is, however, left with a permanently higher price level.

But while this argument applies to labor as a class, it does not necessarily apply to all grades and kinds of labor. Skilled labor is likely to benefit from a protective-tariff policy in a situation such as the one which we have been describing. We have assumed that without the tariff policy only agriculture, mining, and lumbering would flourish. None of these industries requires a large amount of skilled labor. In consequence of the small number of skilled jobs to be filled, wages of high-grade labor would very likely be low relatively to the wages of unskilled and semiskilled labor. Under a protective-tariff system manufacturing would develop, and the demand for skilled labor would increase relatively to the demand for common labor. Under these circumstances it is not unlikely that a protective tariff would raise the real wages of the skilled workman. Thus while the total national income is reduced by a tariff policy, it may still be true that under certain circumstances some of the agents of production actually benefit by a tariff policy.

But now let us consider the second possible condition. Assume that the resources and population of the country are such that not only will agriculture, mining, and lumbering flourish but also a great many forms of manufacturing, even though there were no protective tariff. Such in point of fact is the case in the United States at the present time. Among the numerous products that we can produce more cheaply than other countries and which therefore need no protection are automobiles, factory machinery, farm machinery, electrical appliances, sewing machines, tools and hardware, boots and shoes, packing-house products, the coarser textiles, and many others. These are amply sufficient to furnish a large field for skilled labor under free-trade conditions. Under these circumstances even skilled labor will very likely suffer a decline in real wages under a protective tariff policy. The argument in support of this conclusion is precisely the same as the one developed above with respect to common and semiskilled labor.

We therefore conclude that in countries whose resources and people are adapted to develop many types of manufacturing on a free-trade basis, even skilled labor and technical experts may lose from a tariff policy, because, although the protective system will tend to develop certain kinds of manufactures requiring a high degree of skilled and technical knowledge, it is likely on the other hand to hamper other types of manufacturing (requiring an equal degree of skill) which flourish best on the free-trade basis.

Industries injured by a Tariff Policy. It is admitted, of course, that the protected industries benefit from the tariff. These benefits are easily seen and are likely to impress the public mind. What is likely to be overlooked is the fact that while certain industries benefit from the tariff, others are positively injured. This aspect of the matter is not so easily seen, and this explains the public's ready acceptance of the tariff fallacy. In the United States, for example, it is safe to say that many more industries are injured by the tariff than are benefited by it. All exporting industries suffer from the tariff. On the one side, it reduces their foreign markets, since the effect of a high-tariff policy is necessarily to restrict international trade; on the other side, the tariff raises their costs of production. Moreover, all the domestic industries which have no foreign competition and the products of which are therefore not protected are injured by the tariff policy because their costs are raised by the tariff. Money wages (though not real wages) and the prices of the protected goods are higher as a result of the tariff, and so the costs to the nonprotected industries are raised.

Let us enumerate some of the industries that are injured by the tariff. In the following table we have classified them under the two main heads of exporting industries and domestic industries. The list on page 714 is by no means complete, but it illustrates the more important industries that are adversely affected.

It is frequently said that the United States is a splendid example of the beneficial results of a protective tariff, that American prosperity has been built up on the basis of a high-tariff policy. In point of fact it is much nearer the truth to say that the United States is an outstanding example of the benefits of free trade. Our domestic market is the greatest free-trade market in the world. There are no tariff barriers between the states constituting the American union. Free trade between the states has opened up a vast national market, and this has made possible mass production, intensive division of labor, and the use of elaborate machinery. It is the large-scale machine system of production which is to a large extent responsible for our high standard of living and our high wage level. Other factors are our rich and abundant natural resources, the high efficiency of our people, the efficiency of management, and the abundance of capital. Over 90 per cent of our total national production goes into our domestic markets. Our foreign trade is relatively unimportant. Our vast domestic market, with free trade between the states, is an important and fundamental basis of American prosperity.

Leading Industries injured by Protective Tariff in the United States[1]

A. Exporting industries[2]
  1. Agriculture
  2. Meat-packing
  3. Machinery
  4. Automobiles
  5. Petroleum
  6. Coal-mining

B. Domestic industries
  1. Railroad transportation
  2. Municipal utilities
  3. Lake and river transportation
  4. Building industry
  5. Trade (wholesale and retail)
  6. Hotels and restaurants
  7. Bakery and confectionery
  8. Publishing and printing
  9. Educational institutions
  10. Government service

British and American Tariff Policies. Why have British manufacturers favored free trade while American manufacturers have consistently favored a protective tariff? If we examine the problem historically the answer is not difficult to find. When the British manufacturing industries first developed,

---

[1] Compare with B. M. Anderson, Jr., "The Tariff in an Unbalanced World," in the *Chase Economic Bulletin,* November 19, 1923.

[2] Leading Exports from the United States in 1926

	Value	Percentage of Total Exports
Farm products and foodstuffs:		
Cotton	$814,000,000	17.3
Wheat and wheat flour	285,000,000	6.0
Packing-house products	235,000,000	5.0
Tobacco (leaf)	137,000,000	2.9
Fruits and nuts	112,000,000	2.4
Raw materials (nonagricultural):		
Petroleum	554,000,000	11.8
Coal and coke	204,000,000	4.3
Copper (ore and manufactures)	141,000,000	3.0
Lumber	97,000,000	2.1
Manufactured commodities:		
Machinery (electrical, agricultural, and industrial)	400,000,000	8.5
Automobiles	320,000,000	6.8
Iron and steel products	76,000,000	3.7
Cotton goods	75,000,000	2.8

they had no competitors of significance in any part of the world. British manufacturers therefore had nothing to fear so far as the home market was concerned. Moreover, the superiority of England's manufacturing system soon developed a considerable demand for her manufactured products in various parts of the world. Her growing manufactures required the importation of raw materials for her factories and foodstuffs for her growing factory population. Britain's extractive industries, including her agriculture, faced foreign competition, however. British agriculture already had a tariff system which protected farm products. Under these circumstances the agricultural interests very naturally demanded a continuance of protection, while the manufacturers demanded free trade. In this struggle the manufacturing interests won. The victory injured the landlords, but the net effect was favorable to the general welfare of the people as a whole. As is almost always true, a program of economic reform is likely to injure certain classes, even though the general good is furthered thereby.

The American situation was quite the opposite. Here the agricultural and extractive industries exported a large part of their products and had nothing to fear from foreign competition. The manufactures, however, faced foreign competition in the domestic market. Thus the manufacturers demanded protection, while the interests of agriculture and the extractive industries were linked with free trade. So long as the agricultural South dominated the government, the tariff was accordingly used chiefly for revenue purposes and not primarily for protection. But during the Civil War the manufacturing North acquired control of the government and has retained this control, for the most part, down to the present time. The farmers of the North — the new West — were linked with the manufacturing East in their opposition to the slave aristocracy of the South. Moreover, the farmers and the wage-earners of the North were adroitly won to the protective-tariff policy by the home-market and high-wages arguments.

It was probably inevitable in a century in which inventions, machine processes, and the factory system were rapidly developing that the rising manufacturing class should achieve a dominating position in industry and politics. In point of fact they became the dominating influence in England during the decade of the thirties and in the United States during the decade of the sixties. In England the interests of the manufacturing class were favorable to free trade; in the United States they were favorable to protection. Hence the divergent tariff policies of the two countries.

But the conditions are rapidly changing in the United States. Many of our manufacturing industries now not only supply the domestic market but export large quantities besides. They are able to compete in the markets of the world, and gain nothing from the protective policy. In fact, most of our manufacturing industries have already reached this stage. On the other hand, the time is not very far distant when agricultural products will have to be imported to supply our rapidly growing population. We are in a transition stage. In the future it is agriculture and not manufacturing that will demand protection. But when that time comes it is doubtful whether the agricultural population will be sufficiently strong politically to secure a protective tariff for their products. The interests of the huge wage-earning population and of the manufacturers will be against them. It is to be expected that the United States will eventually go the way of England and become a free-trade country.

But tradition is a powerful force. It will take a long time to change ingrained habits of thought. It is possible that the bulk of our manufacturing industries are even now injured by the tariff. Yet American manufacturers still stand solidly back of the protective principle. The American public has long been schooled in the doctrines of protection, and it will take a long time to change the established tradition. There are, however, already signs that the large banks, in close touch with international trade, are beginning to favor freer trade.

STATEMENTS AND PROBLEMS FOR DISCUSSION

1. Referring to the future revival of European industry and the program of tariff revision, Dr. B. M. Anderson, Jr., of the Chase National Bank, New York, writes as follows: "This would involve a shifting of our industries: Those lines of manufacturing production, in particular, which use a great deal of labor in comparison with the capital which they employ, will probably be obliged to give up part of their labor and to contract their operations. . . . We shall succeed best in those lines where labor is economized, and land and capital used more lavishly. Labor will be released in the United States from some lines of manufacturing, will shift back to agriculture, to the building trades, to copper mining, and to other lines which will be stimulated by the restoration of Europe or by a lower level of costs. This process should, however, involve no general slump, and should rather be accompanied by a general stimulus to business, based on the revival of certain lines which have been unduly depressed. Among these will be copper, zinc, the packing industry, the hide and leather industry, the fertilizer industry, and the farm implement industry, as well as the growing of grain and live stock." Why should we, under conditions of free trade, go in for industries that employ relatively more capital and land and less labor?

2. A tariff must necessarily lower the total national real income. But certain factors of production will gain from a tariff policy; others may lose. Thus farmers would gain from free trade, but it is not impossible that skilled labor as a class may lose. Defend each of these statements.

3. "Since free trade would be particularly injurious to those industries which employ much labor and relatively little capital and land, it follows that labor as a factor in production would tend to be squeezed out if the tariff were removed. It would therefore be less significant in the productive process compared with the other factors of production. Landowners and investors would gain from free trade while labor would lose." Evaluate this argument carefully. If it is admitted that labor would be worse off relatively, might it still be better off absolutely? Why?

4. Is there any ground for the conclusion that free trade might have the effect of raising temporarily the rate of interest? If the rate of interest were raised, is it likely that more capital would be accumulated? Why? What would be the effect of this on the total national income? on real wages?

5. "Protective tariffs on manufactured goods do affect the farmers adversely in two ways. On the one hand, they raise the prices of commodities which the farmers must buy. On the other hand, by checking the imports of goods from Europe, they lessen the supply of dollars available for European purchasers of American farm products and so cut under the market for our farm products and lower their prices. The farmer is on strong ground when he demands the reduction of the tariff on manufactured commodities." — B. M. ANDERSON, Jr. Show that this argument is sound. How do you account for the fact that the American farmers have been willing to support the tariff policy?

6. "The tariff builds up a home market for farm products and hence tends to raise the prices of these commodities." Point out the fallacy in this argument.

7. It is true that if Great Britain adopted a protective tariff certain industries would be stimulated, but what is likely to be overlooked is that her export industries would be injured by such a policy. Show why this would be the case.

8. It is clear that if Great Britain adopted a tariff policy her labor, capital, and land would be employed in a different manner from that prevailing under free trade. The question is Would her people and resources be employed more favorably or less favorably? What is the answer?

9. Some industries which are protected by the low tariffs are really injured more than they are helped by reason of the additions to their costs made by tariffs benefiting other people. The steel industry is probably in this class.

10. It is argued that the chart on page 702 does not in fact show a progress toward free trade but a progress toward higher protection, because they show that comparatively little of dutiable goods was imported, since the duties on dutiable goods are so high as almost wholly to shut off imports. Criticize this argument.

INDEX

Addyston Pipe and Steel Company, 268
Advertising, 88, 159; a form of price competition, 174–176
Aftalion, Albert, on income theory of money, 365
Agents of production, combination of, 50–78; what determines cost of, 163–164; effect of mobility of, on price, 200; marginal use of, 203–204; best combination of, affected by change in demand, 238; better combination of, a cause of decreasing cost, 238–241
Agriculture, cost of final products of, to consumer, 92, 229; effect of price fluctuations on, 341–343, 555–563; movement of land values, 556; rental value of land, 557–558; farm prices in relation to other prices, 557–561; overcapitalization of land, 562–563
Amalgamated Clothing Workers of America, 518
American Federation of Labor, fluctuation of membership with prices, 347; effect of, on immigration legislation, 513; structure of, 517; policies of, 517; growth of, 518–519; on compulsory arbitration, 535
American Smelting and Refining Company, 280
American Sugar Refining Company, 273–274
American Tobacco Company, 270–271, 275–277
Anaconda Copper Company, 280
Anarchism, 610
Anderson, B. M. Jr., on tariff, 714, 717
Angell, J. W., cited, 674
Assumption-of-risk doctrine, 537
Australia, compulsory arbitration in, 535

Balance of payments for 1927, 666
Balance of trade, 661–680; factors governing, 670–671
Bank acceptances, 651
Bank credit, 327–330
Bank notes, 319–321
Banking theory of business cycles, 387–388
Bernstein, Eduard, 609
Bill of exchange, 329
Bimetallism, 322–327
Blanc, Louis, 603
Bland-Allison Act, 325
Böhm-Bawerk, E. von, 609
Bolshevism, 611–614
Bowley, A. L., on income distribution, 596
Boycott, use of in labor disputes, 523
British Unemployment Insurance Act, 543
Business cycles, 373–400; effect of, on costs, 170; theories of, 382–393; relation to unemployment, 521–522. *See also* Stabilization
Business organization, forms of, 43–50
Business profits, 563–568

Canadian Industrial Disputes Act, 534–535
Capital, defined, 464–466; producer's, 466–468; marginal productivity of, 468–470, 479; superior productivity of, 470–472; accumulation of, 573–574
Cassel, Gustav, on stabilization of the value of gold, 370; on business cycles, 391–393; on foreign exchange, 656
Causal relations in economics, 7–13
Chambers, C. R., cited, 558
Child-labor laws, 34–36, 535–537
Clark, F. E., on marketing functions, 87
Clayton Anti-Trust Act, 272–273
Closed-shop controversy, 525–528
Coinage, 318–319, 322–327
Collective bargaining, 524–526
Combination of agents of production, 52–78
Combinations, reason for failure of, 259; not achieving complete monopoly, 261–262; history of, 265–281

Commodity theory of money, 358–365
Common ownership of stock, 267
Communism, 611–614
Communist Manifesto, 605–608
Comparative cost, law of, 681–698
Competition, limited, 260–262, 283–285, 505–506; foreign, 694–698
Compulsory arbitration, 532–535
Concentration, in industry, *see* Large-scale production
Concentration of wealth and income, *see* Inequalities of property-ownership and income
Consolidations, 267
Conspiracy, legal status of, in labor disputes, 528–532
Constant cost, 211–218; defined, 214; conditions giving rise to, 215–217; shifting of costs and, 217–218
Consumers' demand and utility, 140–161
Contributory negligence, 537
Control over utilities, 25–27
Coöperation, 47–48, 621
Copper Export Trading Company, 280
Corporation, 45–47
Cost, lowest unit, 65–72; prime and supplementary, 167–172; normal price and cost of different entrepreneurs, 176–183; overhead distinguished from joint, 206–208; effect of changes in technique on, 244–245; historical trend of, 244–246. *See also* Normal price
Cost of production, relation to price, 163–167
Cost per unit of output, 64–72
Credit instruments, 328–329
Crises, *see* Business cycles
Customs duties, *see* Tariff

Damage suits in labor disputes, 530–531
Dawes plan, 676–680
Day, E. E., cited, 362; on increase in production, 550
Debs, E. V., injunction case, 269
Debt settlement, application of theory of marginal utility to foreign, 159–160. *See also* International debts
Decreasing cost, meaning of, 235; causes of, 237–243; example of, in railroads, 238–240; restricted competition and, 246–248; incidence of new unit costs and, 248–249; compared with increasing cost, 249–251
Demand, defined, 102; individual, 102; collective, 107–108; schedules of, 103–104, 108; elastic and inelastic, 109–112; changes in, 112; effect on market prices of changes in, 131, 211; analysis of, 140–143; relation of psychology to analysis of, 141–142; relation of, to marginal utility, 151–157; effect of decrease in, on earnings of fixed capital, 169–170; effect of changes in, on costs, 212–213.
Demand price, 103
Deposit currency, 327–330
Dewing, A. S., on efficiency of combinations, 273
Diminishing returns, law of, 56–60; a check to expansion of efficient firms, 67–72, 77–78, 187; relation to increasing cost, 223, 230; in relation to growth of population, 251; in relation to rent, 447, 454–456, 557
Diminishing utility, 146–149
Discount policy of Federal Reserve Banks, 396–398
Discount rates, relation to balance of trade, 671. *See also* Interest
Distribution of estates, in Massachusetts, 577–584; in the United States, 582–584; in Minnesota, 584–585
Distribution of income, personal, 404–406, 576–577, 586–600; functional, 403–427, 429–462, 464–491, 493–508, 598–600
Distribution of wealth among living persons, 585–586. *See also* Distribution of estates
Division of labor, 17–20, 215
Documentary bills, 660
Domestic system, 509–510
Douglas, P. H., on wages, 544; on wholesale and retail prices, 546
Dumping, 283
Dynamic aspects of distribution, 544–574

Earned income, 642
Economics, subject matter of, 3–7; as a science, 7–13
Economic goods, 21–22
Economic laws and tendencies, 9–13

Economic order, the present, 30–31
Economic utilities, 21–22
Effective utility, 149–150
Elasticity of currency, 333
Elasticity of demand, 109–112; effect of, on price and gross receipts, 126–127; monopoly control and, 262
Employers' associations, 520
Engels, Frederick, 605–608
Entrepreneur, 23, 40–50; marginal, 65–67; superior, 67–72; work of, described, 495–496; as risk-taker, 501–502
Entrepreneurs' differential costs, 174–188; permanence of, 183–188; increasing cost and, 226–227
Equalitarian distribution considered, 590–592
Equation of exchange, 351–355
Erdman Act, 532
Evolutionary socialism, 608–609
Exchange rates, 654–660
Expenses of production, and price, 162–172, 174–188, 190–218, 220–233, 235–251; and diminishing returns, 224–226
Exploitation, gains of, 506–507

Fabian socialists, 608
Factors of production, 53–55; scarcity of factors, 55–56; combination of, in agriculture, 55–60, 241
Factory legislation, 34–37
Family-allowance wage, 513–514
Farm prices, 555–563
Favorable balance of trade, 663–664
Federal Reserve system, 331–334
Federal Trade Commission, 272–273
Fellow-servant doctrine, 537
Filene, E. A., on minimum wage, 516
Finance bills, 660
Fisher, Irving, on stabilizing the value of money, 367–369; on business cycles, 390–391; on stabilizing business, 395–396; on time preference, 474
Foreign exchange, 649–660; rates of, 654–660
Forms of combination, 265–268
Fourier, Charles, 604
Free-trade principles, 700–716
Friday, David, cited, 567
Functions of money, 317–318
Futures, dealings in, 309

General Electric Company, 279, 281
General Motors Company, 77
Gentlemen's agreements, 267
Gold, as standard of value, 318, 322–327; stabilization of the value of, 369–371; future production of, 370
Gold-shipping points, 654–655
Goodwill, 26–27
Government intervention in labor disputes, 532–535
Governmental enterprises, 48–50
Greenbacks, 319, 324
Gresham's law, 323
Guild socialism, 610–611

Harding, President, on trade unions, 529
Harvard Economic Society, cited, 568
Hawtrey, R. G., on business cycles, 387–388
Health insurance, 537–538
Hedging, 313
High-wages argument for tariff, 705–706, 709–712
Hobson, John A., on business cycles, 383–387; on minimum wage, 516
Holding companies, 267
Home-market argument for tariff, 704–705, 707–709
Hours of labor, shorter work day, 521; legislation concerning, 536–537
Human behavior in economic matters, 8–9

Immigration, 553–555
Imports free of duty, 702–704
Improvements on land, return of, similar to rent, 444
Incidence of taxes, 627–631
Income, capitalization of, 484–486; earned and unearned, 642; tax on individual, 645
Income theory of money, 365–367
Increasing cost, 220–233; meaning of, 220; normal price and, 222; conditions leading to, 222–224; forces counteracting, 227–230; relation of, to diminishing returns, 223, 230; incidence of new unit costs under, 230–231; decreasing cost compared with, 249–251
Index numbers, 336–339
Individualistic order, 30–33

Individualistic philosophy, 31–33
Infant-industry argument for tariff, 703–704, 706
Industrial Disputes Act, Canadian, 534–535
Industrial fluctuations, 373–375. *See also* Business cycles
Industrial revolution, 30, 34, 509–511
Inequalities of property-ownership and income, 405–406, 576–600
Injunctions, 531–532
Installment buying, 357
Institutional economics, 6–7, 52–53. *See also* Book IV, Part II, pp. 509–647
Institutions in present order, 30–31
Interdependent prices, 190–210; principal types of, 190–193
Interest, 464–491; rate of, 482–483, 489–491; types of, 486–487; explicit, 486; implicit, 486; loan market and rate of, 488; effect of supply of money on, 489–491; effect of changes in the price level on, 490–491, 568–573; trend of, 568–574
Interlocking directorates, 267
International concerns, 280–281
International debts, 661–680; agreements with respect to, 674–676
International Harvester Company, 277–278
International Kartells, 279–280
International Match Corporation, 280
Inventions, effect of, on efficiency of labor, 418, 547; effect of, in counteracting diminishing returns, 456

Jaurès, Jean, 609
Jevons, William S., on business cycles, 382–383
Johnson, Judge A. W., cited, 584
Joint Commission of Agricultural Inquiry on cost of marketing, 20, 92
Joint cost, 205–206
Joint demand, 191–192; price and, 197–202
Joint supply, 193; price and, 205–206
Joint Traffic Association, 269
Jones, Elliot, on the trust problem, 273

Kansas Industrial Court, 534
Kartells, 266, 279–280
Kemmerer, E. W., on the stabilization of the value of gold, 370–371
Kennicott Copper Company, 280
Keynes, J. M., on foreign exchange, 656
King, W. I., on incomes, 93; on unemployment, 543; on distribution of wealth and income, 583, 585–586; on annual savings, 591; on functional shares in distribution, 598–599
Knights of Labor, effect of, on immigration legislation, 513; history of, 517

Labor, 509–555, 602–621; earnings of, 406–407; determination of marginal productivity of, 409–415; supply of, 415–423; effect on wages of, differences in supplies of different grades of, 434–437
Laborers, divided into income groups, 429–431
Labor legislation, 532–542; arbitration of labor disputes, 532–535; minimum-wage laws, 535–536; hours of labor, 536–537; social insurance, 537–538; principles of, 538–541
Labor-saving machinery, effect of, on employment, 522–523
Laidler, Harry W., on communism, 613
Laissez-faire, 30–33
Land, use of term in economics, 443; demand for, 446–449; marginal productivity of, 449–450, 455
Land values, 555–563
Large-scale production, 72–78; economies of, decreasing cost and, 240–241
Lassalle, Ferdinand, 603
Layton, W. T., 673
Legal status of trade unions, 528–532
Lehfeldt, R. A., on stabilization of value of gold, 369–370
Liefmann, Robert, on international Kartells, 279
Loan market, 487–489
Lowest unit cost, 65–72

Machine methods and increased productivity of labor, 470–472
McNary-Haugen Bill, 231–233, 234
Malthus, T. R., on free trade, 3; population theory of, 415–416, 420
Marginal buyers and sellers, relation to market price, 132–134
Marginal entrepreneur, 65–67

Marginal product of variable agents, 63–64

Marginal productivity, of labor, 409–415; not final explanation of wages, 423; of land, 449–450, 455; of capital, 468–470, 479

Marginal use of agents of production and price, 203–204

Marginal utility, 149; how related to individual demand, 151–157; market price and, 157–158; practical applications of the theory of, 158–160; distribution of taxation and, 159, 646

Margins, extensive and intensive, 452

Market, 80–94; concepts of, 80–82; producers' and consumers' contrasted, 83–86; related markets, 85; ideal, 176

Market price, 123–138; without reservation of supply, 123–129; with reservation of supply, 129–130; effects of changes of demand and supply on, 130–132; equilibrium and actual contrasted, 134–138; how affected by combinations, 135–136; the theory of marginal utility and, 157–158

Marketing, 80–94; functions of, 86–90; specialization and, 90–94; retail cost of, 91–92; incomes from, 92–94. *See also* Market

Marshall, Alfred, 495, 505

Marx, Karl, 603, 605–608

Marxian socialism, 605–608

Maxwell, W. F., cited, 362

Mediation and conciliation, 532–534

Mergers, 267

Miami Copper Company, 280

Mill, John Stuart, on law of diminishing returns, 56; on monetary theory, 355–356

Minimum-wage legislation, 440–442, 535–536, 538–541

Mitchell, Wesley C., on business cycles, 388–390; cited, 591

Monetary system of the United States, 333–334

Monetary theory, 349–371

Money, kinds of, 318–322; quantity theory of, 355–365; stabilization of the value of, 367–369; supply of, and interest rates, 489–491

Money and credit, 317–334

Monopoly, defined, 253; partial, 254; causes of, 257–260; control affected by elasticity of demand, 262; taxes on profit of, 629–630

Monopoly and restricted competition, 265–285

Monopoly price, 253–263; theory of, 255–257; limitations on, 262–263

Monopoly profit, 256

Moore, Henry L., on yield and price of corn, 12; on business cycles, 382–383

National Bureau of Economic Research, on current income, 638

National-defense argument for tariff, 706, 707

National Industrial Conference Board, on increase of taxation, 636–637

National Labor Union, 517

National Trades Union, 516–517

Natural law in economics, 9–13

Need, as a basis for personal distribution, 599–600

Newlands Act, 532–533

New Zealand, compulsory arbitration in, 535

Nontax revenue, 625–626

Normal price, 162–251; relation of cost of production to, 163, 165–167; under static conditions, 177; under dynamic conditions, 178–181; summary of theory of, 208–210; affected by changes in technique, 244–245. *See also* Price

Normal value, *see* Normal Price

Old-age pensions, 537–538

Open-market operations of Federal Reserve Banks, 398–399

Open-price associations, 267, 272

Open-shop controversy, 525–528

Opportunity cost, 164, 172

Organization of production, 40–50

Organized economy, 283–285

Origin of trade unions, 509–511

Overhead cost, distinguished from joint cost, 206–208

Overproduction, 383–387

Owen, Robert, 604

Panaceas for the labor problem, 541–542

Par of exchange, 654

Partnership, 44
Patents, 258, 278–279
Pegging the rate of exchange, 660
Permanence in the distribution of income, 593–597
Pierson, Dr. N., on economics as a science, 7
"Pittsburgh-plus," 282
Political units as entrepreneurs, 48–50
Politics, and labor, 523–524
Pools, 265–266
Population, growth of, in relation to diminishing returns, 251; Malthusian theory of, 415–416, 420; checks on growth of, 416–417; rates of growth of, in different countries, 417–419
Price, economics of, 6–7; defined, 97; analysis of, 98–100; market and normal contrasted, 100–101; related to cost of production, 163–167; competition and, 174–176
Price fluctuations, 336–347; history of, 339–341; effects of, 341–347; relation to business cycles, 390–391; effect on wages, 548–553
Price policies, 281–283
Prime cost, 167–172
Production, 15–28; superiority of roundabout method in, 470–472
Productive activities, 17–21
Productivity and wages, 546–548
Profits, 493–508; "pure profits" distinguished from "business profits," 493–494; functional distribution and, 495; from earnings of rare ability, 498–499; relation of, to risk-taking, 502–505; from monopoly and exploitation, 506–507; business profits, 563–568
Property as a social institution, 597–598
Proprietorship, individual, 43–44
Public regulation, 33–37, 287–303
Public utilities, 287–303
Public works, building of, as an aid to stabilization, 399
Purchasing-power parity, 655–658

Quantity theory of money, 355–365

Railroad Brotherhoods, 518
Railroad Labor Board, 533
Railroads, 287–303; an example of decreasing cost, 238–240; earnings of, 302; consolidation of, 302–303
Rainfall and business cycles, 382–383
Rates of public utilities, regulation of, 291–301
Regulation of production, 30–37
Rent, 443–462; economic, defined, 444, 445; explicit and implicit, 445; relation to diminishing returns, 447, 451, 454–456; amount of, determined by marginal productivity, 449–450; and land-differences, 450–454; social theory of, 454–456; agricultural, 456–457; urban, 457–459; as an unearned income, 459–462; tax on, 460–462, 628; trend of, 555–563
Reparations payments, effect on foreign exchange, 657–658; effect on balance of trade, 671–674; German, 676–680
Reservation price, 120–121, 123, 129–130
Restrictive policies of trade unions, 512–516
Ricardo, David, on free trade, 3; on importance of distribution, 404
Ripley, William Z., on railroad consolidations, 303
Risk, and uncertainty in production, 305–306; methods of meeting, 306–308; assumption by entrepreneurs, 501–502
Risk-taking, speculation and, 309–311; profits and, 502–506; real cost of, 504–505
Rival demand, 192; price and, 202–205
Rival supply, 190–192; price and, 193–197
"Rule of reason" cases, 269–270

Saving, a prerequisite for production, 24–25; motives for, 475–478; involuntary, 476–477; by corporations, 477–478; supply of, 478–480; marginal cost of, 479–480; demand for, 480–484; social, 573–574
Scarcity, conditional to value, 145–146
Seasonal unemployment, 521–522
Secrist, Horace, on related markets, 85
Security issues, Federal control over, 302
Self-generating cycles, 388–390
Self-interest, a factor in social control, 37
Sherman Anti-Trust Law, 268–270

Sherman Silver Purchase Act, 325–326
Shop committees, 524, 527–528
Sidgwick, Henry, on the tariff, 706
Single tax, 460–462
Smith, Adam, on the nature and object of economics, 7; on protection, 707
Snyder, Carl, on velocity of deposits, 363; cited, 563
Social control, 6–7, 33–37
Social income, 27–28
Social insurance, 537–541
Social legislation, in England, 34–36; in the United States, 36–37
Socialism, 602–621; compared with communism, 613–614; compared with capitalism, 614–615; difficulties encountered by, 615–618; value and distribution under, 619–621
Socialist movement, 602–604
Sorokin, P., on American millionaires, 601
Special assessments, 625, 626
Specialization, 215, 251; and marketing, 90–94
Specie parity, 654–655
Speculation, risk-taking and, 309–311; effect on prices, 311–314; rates of foreign exchange and, 658–659
"Spoiling the market," 116–117, 135
Sprague, O. M. W., on stabilizing business, 396–397
Stabilization, of the value of money, 367–369; of the value of gold, 369–371; of business, 394–400
Stamp, Sir Josiah, on annual savings, 591; on distribution of income, 593–596
Standard of value, 322–327
Standard Oil Company, 270–271, 274–277, 280–281
Stewart, W. W., cited, 362
Stock market, speculation on, 310–311, 314; prices on, 564–567
Strikes, increase of, 523
Style goods, prices of, 118
Subsidy, effect of, under increasing cost, 231–233; on agricultural products, 232
Subsistence theory of wages, 415, 420–421
Substitution, 190–191; conditions determining, 194–197; principle of, in distribution, 413–415, 468–469
Sugar Trust, 268, 273–274
Supplementary cost, 167–172
Supply, and demand as regulative forces, 31; defined, 114; schedules of, 114–120; determination of, 120–121, 162–163; and demand, interaction of, 130–132; unreserved, 123–129; reserved, 123, 129–130
Supply price, defined, 115; significance of, 129
Swedish Match Company, an international concern, 280
Syndicalism, 609–610

Taft, W. H., on collective bargaining, 526–527
Tariff, principles of, 700–716; tariff wall, 701; industries injured by, 712–714; British and American tariff policies contrasted, 714–716
Taussig, F. W., cited, 674
Tax, defined, 624–625; direct and indirect, 626; kinds of, 626–627; shifting and incidence of, 627–631; on economic rent, 628; on surpluses, 628; on net incomes, 629; capitalization of, 631–633; on net incomes of individuals (Act of 1928), 645
Taxation, 622–647; relation to distribution of wealth and income, 623; amount of, 633–635; increase in, 635–639; distribution of, 640–644; progressive and proportional, 644–647
Taxes, on output of farms, mines, etc., incidence of, 231
Tax system, characteristics of a good, 640
Technique, diminishing returns offset by changes in, 241–242; kinds of change in, 242; effect on cost and normal price of changes in, 244–245; improvements in, and wages, 423–427; improvements in and rent, 444
Tendencies, economic conclusions as, 9–13
Theory of money, 349–371
Time preference, causes of, 473; definition of, 474; individual differences in, 474; rates of, 474–475
Trade associations, 272
Trade unionism, 509–542; origin of, 509–511; effect on wages, 511–516, 555; growth of, 516–518; extent of

organization of, 518; strongholds of, 519; demands of, 520–523; methods of, 523–526; place in modern economic order, 526–528; legal status of, 528–532
Trans-Missouri Freight Association, 269
Transportation Act of 1920, 301–303
Trusts, 266–267
Twin Cities Milk Producers' Association, 284, 285

Uncertainty, production for the future and, 305; in economic life, 499–501
Underconsumption, 383–387
Unearned income, 642
Unearned increment, 459–462
Unemployment, causes of, 521–523; extent of, 522; insurance against, 523, 537–538, 543
Unfair practices, 270–272, 273–277
Unfavorable balance of trade, 661–663; economic effects of, 667–670
Unionism, *see* Trade unionism
United States Shoe Machinery Company, 278–279
United States Steel Corporation, 270–271
Utility, kinds of, 19–20; demand and, 143–145; marginal, 149; effective distinguished from total, 149–150
Utilities, economic, 21–22; source of, 21
Utopian socialism, 604

Valuation of public utilities, 292–299
Value, defined, 97; of money, 349–351. *See also* Price; Market Price; Normal Price
Victoria, (Australia) minimum-wage boards in, 535

Wage-adjustments, principles of, 427–428
Wage differences, 429–442; apparent, 431–433; proximate cause of, 433–434; fundamental causes of, 434–437; between places, 437–439
Wages, 403–427, 429–442; defined, 406; different from other prices, 407–408; determined by marginal productivity, 409–413; subsistence theory of, 415, 420–421; related to supply of labor, 415–423; affected by improvements, 423–427; of women, 440; effect of trade unions on, 511–516, 555; rise of, 520–521; long-run trend of, 544–546; effect of productivity on, 546–548; effect of price fluctuations on, 548–553; effect of immigration on, 553–555
Wants, 15–17
War loans, effect on balance of trade, 671–674. *See* International debts
War Labor Board, 534
Watson-Parker Act, 533
Wieser, F. von, on income theory of money, 365; cited, 495
Working, Holbrook, cited, 362
Workmen's compensation, 537
Wright, Judge J. W., cited, 584

Young, Allyn A., cited, 362

Date Due

FEB 12			
OCT 19	1977		
PRINTED IN	U. S. A.	CAT.	NO. 23233

"The Robber Barons" Jo

HB 171.5 .G25
Garver, Frederic Benjamin 010101 000
Principles of economics,
0 1163 0226528 9
TRENT UNIVERSITY

HB171.5 .G25
Garver, Frederic Benjamin
Principles of economics.

DATE	ISSUED TO

17663